

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p mode:

Run on: November 12, 2003, 06:43:44 ; Search time 26 Seconds
(without alignments)
941.7133 Million cell updates/sec

Title: US-09-735-251-3

Perfect score: 2302

Sequence: 1 GCGAGAGGCCAGAGAGAA.....TGCCTCCATCTCCCTGTC 1273

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Egapop 6.0 , Egapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn21/USPTO.spool.p/US09735251/tunat.12112003.061636.28155/app.query.fasta_1.1415
-DB=PIR_76 -OPMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPTCI=0 -LOPEXT=0
-UNITS=bits START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09735251 @CGN.1.1.38 @tunat.12112003.061636.28155 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPB=LOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_76:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.5	8-2	373	2 S54583	ubiquitin-like pro
2	165.5	7-2	839	2 T04859	extensin homolog P
3	164.5	7-1	536	2 B84549	probable ubiquitin
4	163.5	7-1	1162	2 JH0557	exo-alpha-sialidas
5	162	7-0	815	2 B56708	extracellular sign
6	159	6-9	528	2 I47141	gastric mucin (clo
7	152	6-6	354	2 T38404	yeast dak2 homolog
8	147	6-4	1188	2 S49315	extensin-like prot
9	146.5	6-4	930	2 D37271	A-alpha 2 4 protei
10	146	6-3	351	2 S50754	hypothetical prote
11	145	6-3	3421	1 WZBBE6	367K tegument prot
12	144.5	6-3	551	2 C84549	probable ubiquitin
13	144.5	6-3	1468	2 A44345	nucleoporin - rat
14	144	6-3	1400	2 T31555	hypothetical prote

15	142.5	6-2	1006	2 T42731	atrophin-1 related
16	142.5	6-2	3530	2 A59266	unconventional myo
17	141.5	6-1	742	2 F84643	hypothetical prote
18	141	6-1	671	2 T02504	hypothetical prote
19	140.5	6-1	1487	1 EDBEE1	immediate-early pr
20	139.5	6-1	1093	2 I38533	AF17 protein - hum
21	139.5	6-1	1844	1 RRPPTM	genome polyprotein
22	139	6-0	2715	2 T13049	eyelid - fruit fly
23	138.5	6-0	1494	2 T14355	protein-tyrosine-p
24	138	6-0	383	2 S32975	gene BCRF2 protein
25	138	6-0	872	2 S33015	hypothetical prote
26	138	6-0	992	2 A31666	hypothetical prote
27	138	6-0	1249	2 T14150	vesicle associated
28	137.5	6-0	1487	1 EDBEF6	155K transcription
29	136.5	5-9	383	2 T46707	proteophosphoglyca
30	136	5-9	1572	2 S45251	SNF2alpha protein
31	134.5	5-8	785	2 A82123	chemotaxis protein
32	134.5	5-8	1125	2 B41206	microtubule-associ
33	134	5-7	753	2 JC2099	glutenin, high mol
34	134	5-7	1647	2 S45252	SNF2beta protein -
35	133.5	5-8	419	2 T29266	glutenin, high mol
36	133.5	5-8	502	2 T18562	hypothetical prote
37	133.5	5-8	574	2 T43556	Wiskott-Aldrich sy
38	133.5	5-8	574	2 T38819	Wiskott-Aldrich sy
39	133.5	5-8	907	2 S86636	hypothetical prote
40	133.5	5-8	1147	2 T42627	ADP-ribosylation f
41	133	5-8	576	2 T36729	probable serine/th
42	133	5-8	744	2 T35192	glutelin ABC trans
43	133	5-6	789	2 A30843	glutelin high mole
44	132.5	5-8	708	2 D96711	hypothetical prote
45	132.5	5-8	1952	2 T48814	hypothetical prote

ALIGNMENTS

RESULT 1

S54583
ubiquitin-like protein DSK2 - Yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YMC21.02; protein YMR276w
C;Species: Saccharomyces cerevisiae
C;Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999
A;Accession: S54583; S59344
R;Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54582
A;Accession: S54583
A;Molecule type: DNA
A;Residues: 1-373 <PEA>
A;Cross-references: EMBL:Z49704; NID:G825540; PIDN:CAA89774.1; PID:G825542; MIPS:YMR276
A;Experimental source: strain AB972
R;Biggins, S.; Rose, M.D.
submitted to the EMBL Data Library, April 1995
A;Description: The yeast ubiquitin-like proteins are involved in spindle pole body dup
A;Reference number: S59343
A;Accession: S59344
A;Molecule type: DNA
A;Residues: 1-108, 'R', 110-295, 'R', 297-373 <BIG>
A;Cross-references: EMBL:L40587; NID:G786149; PIDN:AAB07267.1; PID:G786151
C;Genetics:
A;Gene: SGD:DSK2
A;Cross-references: SGD:S0004889; MIPS:YMR276w
A;Map position: 13R
C;Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology
F;3-77/Domain: ubiquitin homology <UEH>

Alignment Scores:
Pred. No.: 1.64e-05 Length: 373
Score: 189.50 Matches: 85
Percent Similarity: 40.16% Conservative: 64
Best Local Similarity: 22.91% Mismatches: 123
Query Match: 8.23% Indels: 100
Gaps: 2

US-09-735-251-3 (1-1273) x S54583 (1-373):

```
QY 230 CTGATCTACTGTCGCGAGCTAAAGATGACGACACCTTGACTTCTATGGATTCAA 289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
44 LeuIleTyrSerGlyLysIleLeuLysAspGlnThrValGlySerTyrHisIleGln 63
QY 290 CCGCTGCTCACTGTCATGCTGTCGAAAGTCTCGCTGAACTGATGACAAACCG--- 346
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
64 AspGlyHisSerValHisLeuVal---LysSerGlnProLysProGlnThrAlaSerAla 82
QY 346 ----- 346
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
83 AlaGlyAlaAsnAlaThrAlaThrGlyAlaAlaAaGlyThrGlyAlaThrProAsn 102
QY 347 ----- 347
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
103 MetSerSerGlyGlnSerAlaGlyPheAsnProLeuAlaAspLeuThrSerAlaArgTyr 122
QY 380 TTCCCGGTGTTGCACACT----- 406
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
123 AlaGlyTyrLeuAsnMetProSerAlaAspMetPheGlyProAspGlyAlaLeuAsn 142
QY 407 AGCAGCTCTCTTACAGAGGCGGCTTTAAG--- 439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
143 AsnAspSerAsnAsnGlnAspGlnLeuLeuAaGlnMetMetGluAsnProIlePheGlnSer 162
QY 440 ----- 487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
163 GlnMetAsnGluMetLeuSerAsnProGlnMetLeuAspPheMetIleGlnSerAsnPro 182
QY 488 GGCCTCAGCAGT---GACCTATTGCTCTGAGGTCTCCAGGACAAAGACCTCTCTCT 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
183 GlnLeuGlnAlaMetGlyProGlnAlaArgGlnMetLeuGlnSerProMetPheArgLys 202
QY 545 GTCTTCGCTATCCCAATATGTTGATACGTGGTGGCTGCTTACCCAGCCCTCTGCAAT 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
203 MetLeuThrAsnProAspMet----- 664
QY 605 GCATTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GlnSerMetGlnPheAlaArgMetMetAspProAsnAlaGlyMetGlySerIleGlyLys 231
QY 665 TCTTCCCGGAGTGCCTCCAGCTCATACGCGATATCCAGGTGGTGGCTCTGTTTAA 724
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
232 AlaAlaSerAlaPheProAla----- 241
QY 725 GGGCTCTCAGATCAGATGATTTTATATCCTTACACCAAGATCAGATCTCTAGAT 784
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
242 ----- 257
QY 785 ACTCCAGCTCCGCGCCAGCCTGCTCTGTAAGTGGATGCTGTAAGTGGATGCTGTA 844
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
258 ThrGlyAsnAsnAlaGlyThrAsnAlaGlyThrAsnAlaGlyThrAsnAlaGly 273
QY 845 ACCCAGATGAGCTGGCCAGCCGCTTATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
274 ----- 281
QY 905 ACCCAGCTCTGCGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
282 LeuLeuAsnProAlaLeuAsnProPheAlaAsnAlaGlyThrAsnAlaGlyThr 301
QY 965 CAGTCAGGAGCGCC---ATCAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
102 ProAlaPheAspProAlaLeuLeuAsnMetPheGlnProValGln---AlaSer 320
QY 1022 CAGGCTCTCTGCGCCAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
321 GlnAlaGlnAspThrArgProProGlnAlaArgTyrGlnIleGlnLeuArgGlnLeuAsn 340
QY 1082 GACATGGGCTATCCAGGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1141 AspValGlyPhePheAspPheAspAlaAsnValAlaAlaLeuArgSerGlyLysPro 360
```

```
QY 1141 ATCAAGCAGCCCTGGAGCTCATCTTTGCTGGA 1173
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
361 ValGlnGlyAlaLeuAspSerLeuLeuAsnGly 371
```

RESULT 2

T04859

extensin homolog F28A21.80 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999

C:Accession: T04859

R:Bevan, M.; Mueller, M.W.; Muendlein, A.; Pelber, R.; Bancroft, I.; Mewes, H.W.; May

submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15387

A:Accession: T04859

A:Molecule type: DNA

A:Residues: 1-839 <BEV>

A:Cross-references: EMBL:AL035526

A:Experimental source: cultivar Columbia; BAC clone F28A21

C:Genetics:

A:Map position: 4

A:Introns: 623/3

A:Note: F28A21.80

Alignment Scores:

Pred. No.: 0.000873 Length: 839

Score: 165.50 Matches: 97

Percent Similarity: 34.86% Conservative: 25

Best Local Similarity: 27.71% Mismatches: 121

Query Match: 7.19% Indels: 107

DB: 2 Gaps: 14

US-09-735-251-3 (1-1273) x T04859 (1-839)

```
QY 312 TGCAGAACTCTGCGCTG---AACCATGATCAGAACCGAACCTGTGCACAAAGTGGCTG 368
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
337 CysGlnLeuProArgLeuLeuGluAsnPheThrPheSerTyrAsnPheThrGly----- 354
```

```
QY 369 CCATGAGAGAGTCCCGGTGTTCACACACTGCCC-----TGCA 407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
355 -----GluProProValCysLeuGlyLeuProGlyPheAspArgAsnCysLeu 372
```

```
QY 408 GCAGCTCTCTTACAGGAGCGGTCTTTAAGATGCTCAGCAATAGAGTCTCTGATC 467
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
373 ProAlaArgProAlaGlnArgSerProGlyGlnCysAlaAlaPheSerSerLeu-ProPr 392
```

```
QY 468 AGATCATTTGCGCCAC-----CCAG 488
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 oValAspCysGlySerPheGlyCysGlyArgSerThrArgProProValValProSe 412
```

```
QY 489 CCCTCAGCAGTGCACCTTATGCTCTGGGGTCTCCAGGACAGGAGCTCTTCTCTCT 548
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
412 rProProThrThrPro-----SerProGlyGlySerPro----- 423
```

```
QY 549 TCGCTGATCCCAATATGCTTGATAGTTGTTGCTGTCCAGCCCTGTGTAATGCCA 608
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
424 -----ProSerProSerIleSerPr 430
```

```
QY 609 TTGCTCTGTTCTGCACCTCCGTAGCAGGAGTGCCTCCCAATGCTGGAGTCTCTCT 668
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
430 oSerPro-ProIle-----ThrValProSerProProThr-ThrProSer 444
```

```
QY 669 CCGGAGCATCCCTCCAGCTCATACGGGATATGCCAGGTGGCTTCTCTTTCAAGGCG 728
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
445 ProGlyGlySerProProSerProSerIleVal----- 455
```

```
QY 729 TCTCAGATGATGAGATGACTTTCCCAAAACACAGGTCCACACCTTCTAGCAGTACT 788
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
456 -----ProSerProProSerThrThr 462
```

```
QY 789 CCAGTCCCGCCAGCCTCCCTGGGTACAGTGGAGCTGTCGGCCCGCCGATCACC 848
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
463 ProSerProGlySerProProThrSerProThrThrProThr---ProGlyGlySerPro 481
```



```

Query Match: 7.04% Indels: 91
DB: 2 Gaps: 13
US 09 735-251-3 (1-1273) x B56708 (1-815)
QY 311 CTGCGAAGTCTGGCTGAAGCTATAGAAAGCTGACCTGTGACCAAGTGGCTG 370
DB 490 LeuAlaAspGlyProSerAlaProLeuAlaAspGlyProAlaThrAla 509
QY 371 -----ATAGAGAGTCCGGTCTGTAATGCTTACAGAGTCCCTCTTA 421
DB 420 GluGluArgGluArgGluArgGluArgGluArgGluArgGluArgGluArg 529
QY 422 AGGAGAGGCGCTTTAAGATGCTGACCAATAAAGATCTCTTATGAAATATTC 481
DB 530 AspGlu -----LysArgArgGluGluAlaAlaAlaGlyAlaGlyAla 546
QY 442 ACCCAAGGCTCAGCAAGTACGCTATATCTGCTGCTGCTGCTGCTGCTGCT 541
DB 547 SerGlyGlyProSerThrAspProLeuAlaGlyLeuValIleSerAsp 563
QY 542 TGTGCTTGTGCTATCCCAATATGCTATACCTTACG-----TGTGCTCA 589
DB 564 -----AspArgSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 579
QY 590 CACAGCTGCTGTAATGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 649
DB 580 ProAlaLeuThr-----SerValProAlaAlaAlaProAlaAla 591
QY 650 CCGGAGTACTGCTCTCTT-----TGGGAGAGCTGCTGCTGCTGCTGCTGCT 700
DB 572 ProThrProThrProThrProAlaAlaProAlaAlaProAlaAla 607
QY 701 ATCCAGAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 760
DB 608 -----ProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 612
QY 761 ACCAGTTCACACCTCTAGCATTA-----TCCCGAGGCTCCCGCA 802
DB 812 eThrGlyProGlnProGlnSerAlaGlySerThrGlyProAlaProGlnProAla 642
QY 803 GCTCTCTGGGTACAGTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 862
DB 632 sProPro-----PheGlyProAlaProAlaProAlaProAlaProAla 644
QY 863 ACCG-----CTTGGCTGGCGCAAGTACAGAGAGCTGCTGCTGCTGCTGCTGCT 919
DB 644 oProGlyProLeuProLeuProAlaAlaProAlaAlaProAlaAlaProAlaAla 664
QY 920 ACCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
DB 664 AlaGln-----SerLeuValIleProProAlaAlaAlaAlaAlaAla 975
QY 980 ATCACCAGTATCTTTCACCCAAAGCTACAGATGCTCTGAGGCTCTGGGAGCC 1039
DB 675 o-----GlySerSerThrProGlyValLeuProThrPheProProGlyLeu 691
QY 1040 AGCTTCAGAGCAGTGGCAGCCCACTGCAAGAGTACTGATGGCCATCCAGAG 1099
DB 691 oProProAspAlaGlyGlyAlaProGlnSerSerMetSerGlu----- 705
QY 1100 GATGAGCTGAGCTCGCGCTGCGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1159
DB 706 -----SerProAspValAsnLeuValThrGlnGlnLeuSer 717
QY 1160 TCATCTTGTGTGAGGAGGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1219
DB 719 LysSerGlnValGluAspProLeuProAlaPheSerGlyThrProLysGlySerGly 737
QY 1220 AGAGGCTAC 1228
DB 738 AlaGlyTyr 740

```

RESULT 6

```

147141
gastric mucin (clone PGM 2A) - pig (fragment)
C1Species: Sus scrofa domestica (domestic pig)
C1Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 03-Nov-2000
C1Accession: I47141; S55315
C1Turner, B.S.; Bhaskar, K.R.; Hadzopoulos-Cladaras, M.; Specian, R.D.; Lamont, J.T.
Gastroenterology 106, 203, 1994
A1Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a novel
A1Reference number: 147141; MUID:94102478; PMID:7506218
A1Accession: I47141
A1Status: preliminary; translated from GB/EMBL/DBJ
A1Molecule type: mRNA
A1Residues: 1-528 <TTS>
A1Cross-references: EMBL:U0281; NID:9915205; PIDN:AAC48526.1; PID:G915208
R1Turner, B.S.; Bhaskar, K.R.; Hadzopoulos-Cladaras, M.; Specian, R.D.; Lamont, J.T.
Biochem. J. 308, 89-94, 1995
A1Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
A1Reference number: S55315; MUID:95275264; PMID:7755593
A1Accession: S55315
A1Status: preliminary
A1Molecule type: mRNA
A1Residues: 1-528 <TTS>
A1Cross-references: 38:U0281; NID:9915205; PIDN:AAC48526.1; PID:G915208
C1Superfamily: pig submaxillary mucin

```

```

Alignment Scores:
Pred. No. Length: 528
Score: 159.00 Matches: 91
Percent Similarity: 37.02% Conservative: 53
Best Local Similarity: 23.39% Mismatches: 132
Query Match: 6.91% Indels: 115
DB: 2 Gaps: 15
US 09 735-251-3 (1-1273) x 147141 (1-528)
QY 378 AGTTCGGGTGTGCACTGCTGCGACA-----GCA 410
DB 24 SerSerSerValProLeuProSerThrThrSerValGlnProSerSerGlySer 43
QY 411 GTCTCTTACAGGAGCGGCTTTAAGATGCTCAGCAATAAGAGTCTCTGGATCAGA 470
DB 44 AlaProThrThrSerAlaThrSerValGlnThrSerSerSerSer----- 59
QY 471 TCATTGTGCGCCACCCAGCGCTCAGCAGTACCCCTATTGCTCTTGGGGTTCTCCAGGACA 530
DB 60 -----ProProLeuSerSerThrIleSerValGlnThrSerSerSerSerVal 76
QY 531 AGGAGCTCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
DB 77 ProThrThrSerThrThrSerValGlnPro-----Ser 87
QY 591 CAGCCCTGTCAATGCCATTGTCCTG-----GTTCTGCACTCCGTAGCAGGAGT 640
DB 88 SerSerSerSerAlaProThrThrArgAlaThrSerValGlnSerSerSerSer 107
QY 641 GCCCAATGCCCTGGGACT-----CACTCTCTTCCCGAGCATGCCCTCCAGC 688
DB 108 AlaProIleSerSerThrThrSerValGlnProSerSerSerGlySerValProThrThr 127
QY 689 TCATCCCGGATATGCCAGGTGGCTTCTGTTGAAGGGCTCTCAGATGATGAGATGAC 748
DB 128 SerAlaThrSerValGlnSerSerSerSerAlaProThrThrSerAlaThrSer 147
QY 749 TTTACCCCAACACAGGTCACACCC----- 775
DB 148 ValGlnProSerSerSerSerSerProIleSerSerThrValSerValGlnProSer 167
QY 776 ---TCTAGCACTCTCCCGAGCTCCCGCCAGCTCCCTCCCGGTACAGTGGAGCTCTGG 832
DB 168 SerSerSerAlaProThrThrSerAlaThrSerValGlnProSerSerSerSer 187
QY 833 CCC----- 835

```


Db 298 AlaSerProProAlaProAlaGlnAspThrArgProProGluLeuArgTyrA:AGluGln 317
 QY 1057 CTCAGCAGCTACGTGACATGGGATCCAGAGGATGAGCTAGCCTGGC-GCCTTGCAG 1125
 Db 318 LeuSerGlnLeuAsnGluMetGlyPheValAspPheGluArgAsnValGlnA:AlaLeuArg 337
 QY 1126 GCACCGGTGGGACATCCCAAGCAGCCTGGAGCTCATCTTTGCT 1170
 Db 338 ArgSerGlyGlyAsnValGlnGlyA:AlaGluSerLeuLeuSer 352

RESULT 8
 S49915
 extensin-like protein - maize
 C:Species: Zea mays (maize)
 C:Date: 05-Mar-1995 #sequence_revision 12 May-1995 #text_change 29-Oct-1999
 C:Accession: S49915
 R:Rubinstein, A.L.; Broadwater, A.H.; Lowrey, K.; Bedinger, P.A.
 submitted to the EMBL Data Library, June 1994
 A:Description: Pex genes: pollen-specific genes with extensin-like domains.
 A:Reference number: S49915
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-1108 -RUB>
 A:Cross-references: EMBL:Z34465; NID:9600117; PDB:1C84230.1; PDB:9600118

Alignment Scores:
 Pred. No.: 0.0181 Length: 1188
 Score: 147.00 Matches: 111
 Percent Similarity: 34.33% Conservativity: 36
 Best Local Similarity: 25.58% Mismatches: 270
 Query Match: 6.33% Indels: 115
 DB: 2 Gaps: 20

US-09-735-251-3 (1-1273) x S49915 (1-1188)

QY 98 CCAAGTCTATTCTTCGGTGGCCAGACAGAACTGGAGAACTACTGCTAGAGGGGTAT 157
 Db 262 ProLysSerIleGlyArgMetValGlyThr:LeuAspGlu 274
 QY 158 AGTATTTTCATTCTGAAGCAGCTTATGCTGGCAAACTCCAGAGAGTCTGTCTCCAGACCT 217
 Db 275 ---IleIlePheLeuAsnLysLeuAspGlyCysLeu 287
 QY 218 GAGCTGATGATCTATCTACTGCTGGTGGAACTAAAGATGACCAAGTACTGCTG 277
 Db 288 LeuGluMetGlyLeu 301
 QY 278 TATGGCATTCACCTGGGTGACCTATGATGCTGTGAAATCTGGCTGAACTGAT 337
 Db 302 SerGlyAsnMetLeuValGlyThre 310
 QY 338 CAGAAACCGGAACCTGGGACAAAGTGGCTGCATGAGAGAGTTCCGGGTGTTCACACT 397
 Db 311 ---ProGluGlnLeuSerAsnIleAlaLysLeuGlnLeuAspVal 395
 QY 398 GCCCTGCACAGCAGCTCTCTTACAGGAGAGCGGCTTTAAGATGCTCAAGCAATAAGGAG 457
 Db 326 ---SerArgAsnValPheThrGly:ValH:sgLysSerIleCysGluLeuPro 342
 QY 458 TCTCTGATCATCATTTGGCCACCCAGGCTCAGAGTACCCCTATTGCTCTTGGG 517
 Db 343 AlaLeuValAsnPheSerPheAlaPheAsnPheAsnSerG:u---A:alaValCys 361
 QY 518 GTTCTCCAGGACAGGACCT 577
 Db 362 MetProSerAspLysAlaLeuVal:AsnLeuAspAspArgAspAsnCysLeuGlyAlaLeu 381
 QY 578 GTGCTGCTGCAC 616
 Db 392 ArgProAlaGlnLysThrAlaLeuGlnCysAlaProValLeuAlaArgProValAspCys 401

QY 617 GTTCTGCACTCGTAGCAGGC-----AGTGCCCAATG 649
 Db 402 SerLysHisValCysAlaGlyTyrProThrProGlyGlyGlyProProSerSerProVal 421
 QY 650 CCTGGAGCTGACTCTCTCTCCGGAGCATGCGCTCCAGCTCATACCGGATATGCCAGT 709
 Db 422 ProGlyLysProAlaAlaSerAlaProMetProSer----- 433
 QY 710 GCTTCTCTGTTTGAAGAGGCTCTCAGATGATGAGGATGACTTTCACCCAAACACAGGTCC 769
 Db 434 -----ProHisThrProProAspValSerProGlu-ProLeuPr 446
 QY 770 ACACCTCTAGCAGTACTCCAGCTCCCGCCAG----- 803
 Db 446 GlnProSerProVal---ProAlaProAlaProMetProMetProThrProHisSerPr 465
 QY 804 -----CCTCCCTGGGTACAGTGGAGCTGCTGGGCCCCGGCCCATC 844
 Db 465 oProAlaAspAspTyrValProProThrProProVal-----ProGlyLysSe 481
 QY 845 ACCGAGTGAAGTGGCCACCGCTTGGCCCTGGCCAGCACTCCGGAGAGAGCTCTCTAC 904
 Db 481 rProProAlaThrSerProSerProGlnValGlnProPro-----AlaAlaSerTh 498
 QY 905 ACACCGACTCTGGCAGCCAGGCTCATCTCTCAGGAGCCTCACCAGATGCTCTCTG----- 959
 Db 498 rPro-----ProProSerLeuValLysLeuSerProProGlnAlaProValGlySe 515
 QY 960 -----GTGTCAGTCAGGAGCGCCATCCACCAATGATCTTTACGCAAGCC 1006
 Db 515 rProProProProValLysThrSerProProAlaPro---IleGlySer-ProSerP 534
 QY 1007 CTACAGCATGCCCTTCAGGCTCTGGCAGCCAGCCAGCTTCAGAGCAGTGGGAGCCCGC 1066
 Db 534 ro-----ProProProValSerValValSerProProProValLysSerProP 551
 QY 1067 CTGCAGCAGTACGTGACATGGGATCCAGGAGCAGTACGCTGCGGCTCGCAGG 1126
 Db 551 roProProAlaProValGlySerProProProProGlyLysSerProProProAla 570
 QY 1127 CCA-----CCGCTGGGACATCCAGCAGCCCTCGAGCTCATCTTTGCTGGAGGAG 1177
 Db 571 ProValAlaSerProProProProValLysSerProProProThrLeuValAlaSer 590
 QY 1178 CCCATGAACCTCCCTGCTTCCCTGAGCCCGCAGCAAGTTCGAGAGGCTACTGCCCTTGG 1237
 Db 591 ProProProProValLysSerProProProProAlaProValAlaSer-----ProPro 608
 QY 1238 GAGGCACATCATGAAGTGGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1273
 Db 609 ProProValLysSerProProProProThrProVal 620

RESULT 9
 D37271
 A-alpha 2,4 protein - bracket fungus (Schizophyllum commune)
 C:Species: Schizophyllum commune
 C:Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 23-Feb-1997
 C:Accession: D37271
 R:Ullrich, R.C.
 submitted to the Protein Sequence Database, October 1991
 A:Reference number: A37271
 A:Accession: D37271
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-930 <ULL>
 C:Keywords: DNA binding; nucleus; transcription regulation

Alignment Scores:
 Pred. No.: 0.019 Length: 930
 Score: 146.50 Matches: 94
 Percent Similarity: 34.82% Conservativity: 62
 Best Local Similarity: 20.98% Mismatches: 154
 Query Match: 6.36% Indels: 138


```

QY 734 GATGATGAGGATGACTTTTACCCAAACACACGAGTCCACACCTCTTAGC----- 781
Db 1289 -----ThrSerSerSerGlySerPhePheValPhe 1299
QY 782 -----AGTACTCCAGCTCCCGCCGACGAC-----TCC 808
Db 1300 GlyThrG:YHisSerAlaProSerAlaSerProAlaPheGlyAlaAanGlnThrProThr 1319
QY 809 CTGGGGTACAGT---GGAGCTGCTGGGCGCCGCGCCATCACCCAGAGTGAGCTGGCCAC 865
Db 1320 PheGlyGlnSerG:GlyAlaSerGlnProAanProProSerPheGlySerIleSer 1339
QY 866 GCCTGGCGCTCGCCGACG-----ACTCCG 889
Db 1340 SerThrAlaLeuPheSerAlaGlySerGlnProValProProThrPheGlyThrVal 1359
QY 890 GAGAGCAGCTCTCACACACCGACT-----CCTGGCACCCAGGGTCAITTCCTCA 937
Db 1360 SerSerSerSerGlnProProValPheGlyGlnGlnProSerGlnSerAlaPheGlySer 1379
QY 938 GGGACCTTCACCAATGCTCTGTGTCAGTCAGGACGGCCCATCACCAATGATCTCTTC 997
Db 1380 G.YThrAlaAasAlaSerSerValPheGlnPheGlySerSerThrThrAsn-----Phe 1397
QY 998 AGCCAAAGCCCTACACATGCCCTTCAGGCGCTC-----TGGCGACCCAGCCCTTCAGAC 1051
Db 1398 AsnPheThrAsnAsnAsn--ProSerGlyValPheThrPheGlyAlaSerProSer 1415
QY 1052 CAGTGGCAGCCCGACGCTCAGCAGCTACGCTGACATGGGCATCCAGGACGATGAGCTCAGC 1111
Db 1416 -----ThrProAlaAlaAlaGlnProSerGlySerGlyGly-PheSerPheSer 1432
QY 1112 CTGGCGGCTTCAGGCGCCCGCTGGGACATCC---AAGCAGCCCTGGAGCTCATCT 1165
Db 1433 GlnSerProAlaSerPheThrValGlySerAsnGlyLysAanMetPheSerSer 1451

RESULT 14
T31555
hypothetical protein Y53H1C.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
C:Accession: T31555
R:White, S.
submitted to the EMBL Data Library, October 1999
A:Reference number: Z1044
A:Accession: T31555
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1400 <W1>
A:Cross-references: EMBL:AL117201; PIDN:CAB55064.1; CESP:Y53H1C.2
A:Experimental: source: clone Y53H1C
C:Genetics:
A:Gene: CESP:Y53H1C.2
A:Introns: 35/3; 78/2; 125/3; 172/3; 224/3; 234/2; 477/1; 527/1; 630/1; 1205/3; 1268/1

Alignment Scores:
Pred. No.: 0.0299 Length: 1400
Score: 244.00 Matches: 90
Percent Similarity: 38.24% Conservative: 32
Best Local Similarity: 28.21% Mismatches: 100
Query Match: 6.26% Indels: 97
DB: Gaps: 18

US-09-735-251-3 (1-1273) x T31555 (1-1400)
QY 337 TCAGAAACCGAACCTGTGTGGACAAAGTGGCTGCCATCAGAGAGTTCGGGTGTTGCCAC 396
Db 988 AlaGluValGlyAlaGlnGlyProGlyAlaHisPheGlyAlaHisGlyAlaSerAla 1007
QY 397 TGCCTGTGCACACGAGCTCTCTTACAGGAGCGCGTCTTTAAGATGCTCAGCAATAGGA 456
Db 1008 ProProProThrSerTyr----- 1013

```



```
Db      278 oHisProGlnProSerLeuHisGlyGlnGlyProGlyProHis-SerLeuGlnThrG 298
Qy      1004 GCCCT-----ACGCATGCCCTTCAGCCCTC-----TGGCAGGCC 1039
Db      299 lyProLeuLeuGlnHisProGlyProGlnProGlnProGlyLeuThrProGlnSerSerG 318
Qy      1040 AGCCTTCAGAGCCAGTGGCAGCCGCCAGCTCCAGACACTAGCTAGCATGGSCAT----- 1092
Db      318 lngGlnGlyProLeuGlyProSerProAlaAlaA-----HisProHisSerThrI 336
Qy      1093 -----CCAGCAGCATGAGCTGAGCTGCCGCTGCGAGGCCACCGGTGGGACATC 1144
Db      336 leGlnLeuProAlaA-SerGlnSerAlaLeuGlnProGlnGlnProPro-----Arg 352
Qy      1145 AAGCAGCCCTGGAGCTCATCTTTGCTGGAGAGAGCCGCAT----- 1183
Db      353 GluGlnProLeuProProAlaProLeuAlaMeProHisIleCysProProProThrThr 372
Qy      1184 GAACCTGCTGCTCCCTGACCC-----CCAGCAAGTTCAGAGGCTAC 1228
Db      373 ProIleProGlnLeuProAlaProGlnAlaHisLysHisProProHisLeuSerGlyPro 392
Qy      1229 TGCCTTGGGAGGCACTCATGAAGGTGCCCTCCA 1261
Db      393 SerProPheSerMetAsnAlaAsnLeuProPro 403
```

Search completed: November 12, 2003, 06:50:12
Job time : 48 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OK nucleic - protein search, using frame, plus n2p -odel

Run on: November 12, 2003, 06:40:09 / Search time 17 Seconds
(without alignments)
7042.94: Million cell updates/sec

Title: US-09-735-251-3

Perfect score: 2362

Sequence: 1 GCAGAGGGCCAGAGGAGAA..... TGGCTGATCTCCCTGTC 127:

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 2.5
Ygapop 10.0, Ygapext 2.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame, n2p_model -DEV=xlp
-C=/con2.1/USP01/spec p/US09735251/runat.12.12.03.061635.28.12.00-app query.fasta.1.14.5
-DB=USP01/USP01/FASTA -FASTA=FASTA -MINMATCH=0 -LGOPT=0 -LGOEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=DJVM62 TRANS=TRANS -LGO=CGI -LIST=45
-DOALIGN=200 -THR SCORE=500 -THR MIN=0 -ALIGN=15 -MCS=LOCAL
-OUTMT=PTO -NORM=EXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=20000000
-USER=US09735251 -ACGN 1.16 @runat.12.12.03.061635.28.12.00 -NCPJ 6 -ICPU=1
-NO.MMAP -LARGEOUERY -NEG SCORES=0 -WAIT -DSBLOCK=0 -LING=0
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREAS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELCP=6 -DELEX=7

Database : SwissProt_41.1

Pred. No. is the number of results predicted by change to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	188.5	8.2	373	1 DSK2 YEAST	P48510 saccharomyc
2	163.5	7.1	1162	1 TCNA TRYCR	P3253 trypanosoma
3	162	7.0	815	1 MK07 HUMAN	Q13164 homo sapien
4	159	6.9	802	1 ENAH MOUSE	Q31373 mus musculu
5	152	6.6	354	1 YANG SCHRC	Q12169 schizosacch
6	148	6.4	1083	1 T2D3 HUMAN	Q0268 homo sapien
7	145	6.3	806	1 MK37 MOUSE	Q9W58 mus musculu
8	145	6.3	3421	1 TEGU HSVER	P28955 equine herp
9	144.5	6.3	1468	1 N153 RAT	P49791 rattus norv
10	143.5	6.2	555	1 DAB1 HUMAN	Q7553 homo sapien
11	143	6.1	830	1 SREC HUMAN	Q14162 homo sapien
12	142.5	6.2	3530	1 MY15 HUMAN	Q9UK77 homo sapien
13	142	6.2	902	1 NFC4 HUMAN	Q14534 homo sapien
14	141.5	6.1	555	1 DAB1 MACFA	Q9BGX5 macaca fasc
15	141.5	6.1	940	1 MAZ4 SCHCO	P37938 schizophyl
16	140.5	6.1	1487	1 ICP4 HSVER	P17473 equine herp
17	139.5	6.1	1093	1 AF17 HUMAN	P55128 homo sapien
18	139.5	6.1	1844	1 POLR_TYMA	P20123 turnip yell

P03204 epstein-bar
P28925 equine herp
Q92558 homo sapien
P27546 mus musculu
P51532 homo sapien
O60641 homo sapien
Q9NZM4 homo sapien
Q9Y4H2 homo sapien
Q15428 homo sapien
Q9JKA7 rattus norv
Q9JKA7 rattus norv
P51531 homo sapien
P28477 turnip yell
P08640 saccharomyc
P14787 oryctolagus
Q8N3F8 homo sapien
Q9TV66 oryctolagus
Q9J104 mus musculu
P24928 homo sapien
P08775 mus musculu
P19532 homo sapien
Q9Y566 saccharomyc
P38968 saccharomyc
P10358 turnip yell
O62203 mus musculu
P10220 herpes simp
P37198 homo sapien
P51666 canis famil

ALIGNMENTS

RESULT 1

DSK2 YEAST STANDARD: PRT; 373 AA.
ID DSK2 YEAST
AC P48510;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ubiquitin-like protein DSK2.
GN DSK2 OR SHE4 CR YMR276W OR YW8021.02.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96281973; PubMed=8682868;
RA Riggins S., Ivanovska I., Rose M.D.;
RT "Yeast ubiquitin-like genes are involved in duplication of the
microtubule organizing center."
RL J. Cell Biol. 133:1331-1346(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX PubMed=9169872;
RA Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome
XIII."
RL Nature 387:90-93(1997).
CC -!- FUNCTION: INVOLVED, WITH RAD23 IN SPINDLE POLE BODY DUPLICATION.
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: TO S.POMBE SPAC26A3.16.
CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
CC -!- SIMILARITY: Contains 1 UBA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

605 GCACATGCTCTGGTTCTGCACCTCCGTAGCAGGAGTCCGCCAATGCTGGGACTGACTCC 664
Db : : : : :
212 GlnSerMetGlnPheAlaIatMetMetAspProAsnAlaGlyMetGlySerAlaGlyGly 231
Qy : : : : :
665 TCTTCCCGAGCAGTCCCTCCAGCTCATCGGAGATATCCAGGTGCCTTCCTGTTTGA 724
Db : : : : :
212 AlaAlaSerAlaPheProAla : : : : : -ProGlyGly- 241
Qy : : : : :
725 GGGCTCTCAGATGATGAGGATGACTTTCACCCAAACACAGGTCCACACCTCTACAGT 784
Db : : : : :
242 : : : : :
Qy : : : : :
785 ACTCCAGTCTCCGGCCAGCCTCCCTGGGTACAGTGGAGCTGCTGGGCCCGGCCCATC 844
Db : : : : :
258 ThrGlyAsnAsnAlaGlyThrAspAlaGlyThrAsnAlaGlyAlaAsn 273
Qy : : : : :
845 ACCCAGAGTGAAGTGGCCATCCGCTTGGCCCTCGCCAGCACTCCGAGAGCAGCACTCTC 904
Db : : : : :
274 : : : : : -ThrAlaAlaAsnProPheAlaSer- 281
Qy : : : : :
905 ACACGGATCTCTGGCAGCCAGGTCATCTCTCAGGAGCTCACTCAATGCTCTCGGTGC 964
Db : : : : :
282 LeuLeuAsnProAlaLeuAsnProPheAlaAsnAlaGlyAsnArgAlaSerThrGlyMet 301
Qy : : : : :
965 CAGTCAGGAGCGCCC--ATCACCAATGATCTCTCAGCCAAGCCCTACAGCATGCCCT 102
Db : : : : :
302 ProAlaPheAspProAlaLeuLeuAlaSerMetPheGlnProValGln--AlaSer 320
Qy : : : : :
1022 CAGGCTCTGGCAGCCAGCCTTCAGAGCAGTGGCAGCCCGAGCTTCGACAGCAGTACGT 108
Db : : : : :
321 GlnAlaGluAspThrArgProProGluGluArgTyrGluHisGlnLeuArgGlnLeuAsn 340
Qy : : : : :
1082 GACATGGGCATCCAGGAGCATGAGCTGAGCT--CGGCGCCCTGAGGCCACCGGTGGGAC 114
Db : : : : :
341 AspMetGlyPhePheAspPheAspArgAsnValAlaAlaLeuArgSerGlyGlySer 360
Qy : : : : :
1141 ATCCAAAGCAGCCCTGGAGCTCATCTTTCTCTGGA 1173
Db : : : : :
361 ValGlnGlyAlaLeuAspSerLeuLeuAsnGly 371

RESULT 2

TCNA_TRYCR

ID TCNA_TRYCR STANDARD; PRT; 162 AA.

AC P23253;
DT 01-NOV-1991 (Feb. 26, Created)
DI 01-NOV-1991 (Feb. 26, Last sequence update)
DE 28 FEB-2003 (Feb. 41, Last annotation update)
DE S:alidase (EC 3.2.1.18) (Neuraminidase) (NA) (Major surface antigen).
GN TCNA
OS Trypanosoma cruzi.
OC Eukaryota; Euzoenozoa;
OX NCBI_TaxID=5691;
{1}_RN
SEQUENCE FROM N.A.
RC STRAIN=Silvio X-10/4;
RX MEDLINE=91277609; PubMed=1711561;
RA Pereira M.E.A., Mejia J.S., Ortega-Barria E., Matzilevich D.,
RA Prioli R.P.;
RT "The Trypanosoma cruzi neuraminidase contains sequences similar to
RT bacterial neuraminidases, YMD repeats of the low density lipoprotein
RT receptor, and type III modules of fibronectin.";
RT J. Exp. Med. 174:179-191(1991).
RN [2]
SUBCELLULAR LOCATION.
RP MEDLINE=91376547; PubMed=1896773;
RX Prioli R.P., Mejia J.S., Aji T., Aikawa M., Pereira M.E.A.,
RA Prioli R.P., Mejia J.S.;
RT "Trypanosoma cruzi: localization of neuraminidase on the surface of
RT trypomastigotes";
RT Trop. Med. Parasitol. 42:146-150(1991).
CC -!- FUNCTION: DEVELOPMENTALLY REGULATED NEURAMINIDASE IMPLICATED IN
CC PARASITE INVASION OF CELLS.
CC -!- CATALYTIC ACTIVITY: cleavage of alpha-(2->3)-, alpha-(2->6)-,

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBJ: L40587; AACB7267.1; ...
 SMLB: Z49704; CAA89774.1; ...
 PIR: S54583; S54583; ...
 HSP: Q15843; NCD; ...
 SGD: S604889; DSK2; ...
 InterPro: IPR036636; ST11; ...
 InterPro: IPR000449; UBA_domain; ...
 InterPro: IPR00626; Ubiquitin; ...
 Pfam: PF00627; UBA; 1; ...
 Pfam: PF02240; Ubiquitin; 1; ...
 PRINTS: PR00348; UBIQUITIN; ...
 SMART: SM00727; ST11; 2; ...
 SMART: SM00165; UBA; 1; ...
 SMART: SM00213; UBQ; 1; ...
 PROSITE: PS00100; UBA; 1; ...
 PROSITE: PS00299; UBIQUITIN_1; 1; ...
 PROSITE: PS00053; UBIQUITIN_2; 1; ...
 KW Nuclear protein.
 FT DOMAIN 1 76 UBIQUITIN-LIKE.
 FT DOMAIN 127 171 UBA.
 FT CONFLICT 109 109 F -> A (IN REF. 21).
 FT CONFLICT 296 296 P -> A (IN REF. 21).
 SEQUENCE 373 AA; 39516 MW; 25E9F82594B0D66 CRC64;

Alignment Scores:
 Pred. NC.: C.000107 Length: 373
 Score: 188.50 Matches: 85
 Percent Similarity: 40.16% Conservative: 64
 Best Local Similarity: 22.91% Mismatches: 123
 Query Match: 8.13% Indels: 100
 DB: Gaps: 12

US 09-735-251-3 (1:1273) x DSK2 YEAST (1:373)

CY 230 CTGATCTACTGTGTCGGAAGCTAAAGATGATGACAGACACTTGACTTCTATGGCATTTAA 289
 |||||
 Db 44 leu1letySerGyLySilleLeLysAspAspGltThiValGluSerTyHisIleLein 63

CY 290 CTTGGTCCATGCTCCATGTCGCGAAAGTCTACCTGACCTGACCTGATGACAAACG... 346
 |||||
 Db 64 AspGyHisSerValHisLeValLysSerGlnProGlySerProGlnAlaSerAla 82

CY 346 ... 346
 Db 83 AlaGlyValAsnAspAsnAlaSerAlaSerAlaSerAlaSerAlaSerAlaSerAla 102

CY 347 ... 347
 Db 103 MetSerSerGlyGlnSerArgGlyPheAspProLeuAlaSerLeuThrSerAlaArgTy 122

CY 380 TTCCGGGTGTGGACACT... 406
 |||||
 Db 123 AlaGlyTyTyLeuAsnMetProSerAlaAspMetPheGlyProAspGlyTyAlaLeuAsn 142

CY 407 AGCAGCTCTCTTACAGGAGCGGCTCTTTAAG... 439
 |||||
 Db 143 AsnAspSerAsnAsnGlnAspGluLeuLeuArgMetGluAsnProIlePheGlnSer 162

CY 440 -----ATGCTCAGCAATAGAGAGCTCTCTGATCAGATCATTTGTGCCACCCCA 487
 |||||
 Db 163 GlnMetAsnGluMetLeuSerAspProGlnValLeuAspPheMetIleGlnSerAsnPro 182

CY 488 GGCCTCAGCAGT---GACCCTATTGCTCTTGGGTCTCTCCAGCAAGACCTCTTCTCT 544
 |||||
 Db 183 GlnLeuGlnAlaMetGlyProGlnAlaArgGlnMetLeuGlnSerProMetPheArgGln 202

CY 545 GTCTTCGCTGATCCCAATATGCTTGATACGTTGGTGGCTGCTCACCAGCGCTCTCAAT 604
 |||||
 Db 203 MetLeuThrAsnProAspMet-----ilArg 211

OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN 1.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=9527943; PubMed=7759517;
 RA Zhou G, Bao Z-Q, Dixon J E.
 RT "Components of a new human protein kinase signal: transduction
 RT pathway.";
 RJ J. Biol. Chem. 270:12665-12669(1995);

RN 2.
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=95374539; PubMed=7646528;
 RA Lee J-D, Ulevitch R-J, Han J;
 RT "Primary structure of BMK1: a new mammalian map kinase.";
 RL Biochem. Biophys. Res. Commun. 213:715-724(1995).
 CC FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND
 CC NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS.
 CC ENZYME REGULATION: Activated by tyrosine and threonine
 CC phosphorylation (by similarity).
 CC TISSUE SPECIFICITY: EXPRESSED IN MANY ADULT TISSUES. ABUNDANT IN
 CC HEART, PLACENTA, LUNG, KIDNEY AND SKELETAL MUSCLE. NOT DETECTABLE
 CC IN LIVER.
 CC DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN
 CC TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.
 CC PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES. WHEN
 CC THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY
 CC ROLE, IS ABSENT.
 CC SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).

CC EMBL: U25278; AAA81381.1;
 CC EMBL: U29725; AAA82931.1;
 CC EMBL: U29726; AAA82932.1;
 CC EMBL: U29727; AAA82933.1;
 CC PIR: B56708; R56708.
 CC KSSP: P24941; IRLC.
 CC Genew: HGNC:16980; MAPK7.
 CC MIM: 602521;
 CC GO: GO:0004707; F:MAP kinase activity; TAS
 CC GO: GO:0007165; P:signal transduction; TAS.
 CC InterPro: IPR003527; MAP kin.
 CC InterPro: IPR000719; Prot_kinase
 CC InterPro: IPR002290; Ser_thr_kinase.
 CC Pfam: PF00069; pkinase; 1.
 CC ProDom: PD000001; Prot_kinase; 1.
 CC SMART: SM00220; S_TK; 1.
 CC PROSITE: PS01351; MAPK; 1.
 CC PROSITE: PS00107; PROTEIN KINASE ATP; 1.
 CC PROSITE: PS00107; PROTEIN KINASE_DXX; 1.
 CC PROSITE: PS00109; PROTEIN KINASE_ST; 1.
 CC Transferrase; Serine/threonine-protein kinase. ATP binding. Cell cycle;
 CC Phosphorylation.
 CC FT DOMAIN 54 346 PROTEIN KINASE.
 CC FT DOMAIN 337 340 POLY-ALA
 CC FT DOMAIN 433 464 PRO-RICH 1.
 CC FT DOMAIN 520 523 POLY-ARG.
 CC FT DOMAIN 577 699 PRO-RICH 2.
 CC NP_BIND 60 68 ATP (BY SIMILARITY).
 CC BINDING 83 83 ATP (BY SIMILARITY).
 CC ACT_SITE 181 181 BY SIMILARITY.
 CC MOD_RES 218 218 PHOSPHORYLATION (ACTIVATES THE KINASE)

FT MOD_RES 220 220 (BY SIMILARITY).
 FT PHOSPHORYLATION (ACTIVATES THE KINASE)
 FT (BY SIMILARITY).
 FT AREGTRPHRCLCS -> GPVVEPANTAASVA (IN
 FT REF. 2).
 FT CONFLICT 609 609 L -> V (IN REF. 2).
 SQ SEQUENCE 815 AA; 88636 MW; 379AD69803207CCF CRC64;
 Alignment Scores:
 Pred. No.: 0.00555 Length: 815
 Score: 182.00 Matches: 89
 Percent Similarity: 35.69% Conservative: 217
 Best Local Similarity: 27.38% Mismatches: 118
 Query Match: 7.04% Indels: 91
 DB: Gaps: 13
 US-09-735-251-3 (1-1273) x MKC7_HUMAN (1-815)
 QY 311 CTCGAAAGTCTCTGGCTCAACCTGATCAGAACCGAACCTGTGGCAAAAGTGGCTGCC 370
 DB 490 LeuArgAspGlyProSerAlaProLeuGluAlaProGluProArgLysProValThrAla 509
 QY 371 -----ATGAGAGAGTTCGGGTGTTGACACTGCCCTGCACAGCAGCTCTCTTAC 421
 DB 510 GlnGluArgGluArgGluGluLysArgArgGlnGluArgLysGlu 529
 QY 422 AGGAGGCGGTCTTAAAGATGCTCAGCAATAAGAGTCTCTGGATCAGATCATTTGGGCC 481
 DB 530 ArgGlu-----LysArgArgGlnGluArgGluArgGluArgGluArgGluAla 546
 QY 482 ACCCGAGGCTCAGCAGTACCTATTGCTTCTGGGTCTCTCCAGGACAGGAGCTTCTTC 541
 DB 547 SerGlyGlyProSerThrAspProLeuAlaGlyLeuValLeuSerAspAsn----- 563
 QY 542 TCTGTCTTCGCTGATCCCAATATGCTTGATACGTTGGTG-----CCTGCTCAC 589
 DB 564 -----AspArgSerLeuLeuGluArgTrpThrArgMetAlaArgProAlaAla 579
 QY 590 CCAGCCCTGTCATCAATGCCATTGCTGTTGCTGCTACCTCCGTAGCAGGAGTGCCTCAATG 649
 DB 580 ProAlaLeuThr-----SerValProAlaProAlaProAla 591
 QY 650 CTGGGAGTGAATCTCTT-----CCCGAGAGTCCCTCCAGCTCATACCGGAT 700
 DB 592 ProThrPro ThrProThrProValGlnProThrSerProProGly----- 607
 QY 701 ATGCCAGTGGCTCTCTGTTGAAGGGCTCTCAGATGATGAGATGATTTTACCCCAAC 760
 DB 608 -----ProLeuAlaGlnPr 612
 QY 761 ACCAGTTCACACCTCTTAGCAGTA-----CTCCAGCTCCCGCCCA 802
 DB 612 oThrGlyProGlnProGlnSerAlaGlySerThrSerGlyProValProGlnProAlaCy 632
 QY 803 GCCTCCCTGGGTACAGTGAGCTGCTGGGCGGCGCCCATCCAGCAGTGGCTGGCC 862
 DB 632 sProPro-----ProGlyProAlaProHisProThrGlyPr 644
 QY 863 ACCG---CCTTGGCTGGCCAGACCTCCGAGAGCAGCTCTCACACACCGACTCTCTGGC 919
 DB 644 oProGlyProLeuProValProAlaProProGlnIleAlaThrSerThrSerLeuLeuAl 664
 QY 920 ACCCAGGTCATTCTCAGGAGCTCACCAGTCCCAATGTCTCTGTGTTCAGTTCAGGAGCCGC 979
 DB 664 aAlaGln-----SerLeuValProProProGlyLeuPr 675
 QY 980 ATACCAATGATCTCTTACCCAGGCTTACAGCATGCCCTCAGGCTCTTGGGAGGCC 1039
 DB 675 o-----GlySerSerThrProGlyValLeuProThrPheProProGlyLeuPr 691
 QY 1040 AGCCTTCAGAGCCAGTGGAGCCCGAGCTGCACAGCTAGTCAGTCATGGCATGCCAGGAC 1099
 DB 691 oProProAspAlaGlyGlyAlaProGlnSerSer--MetSerGlu----- 705

QY 278 TATGGCAATCAACCTGGTCCACTGTCCATGTCTCGAAAGTCCTGGCTGAACCTGAT 337
 DQ 308 -----ProProThrSer 311
 QY 338 CAGAAACCGAAGCTGTGGCAAAAGTGGTGCATGAGAGAGTTCGGGGTGTGCACACT 397
 DQ 312 ThrProProThrProProLeuArgHisA:aaIaThrArgPheAlaThrSerLeuGlySer 331
 QY 398 GCCTGTCAGCAGGAGCTCTTACAGGGAGGCGTTCAGATCTCTCAGCAATAGGAG 457
 DQ 332 AlaPheHisProValLeuProHisTyrAlaThrValProArgProLeuAsnLysAsn 350
 QY 458 TCTCTGGATCAGATCATTTGGCCACCCCGACGCTCAGCAGTGCACCTATTGCTCTGGG 517
 DQ 351 SerArgProSerSerProValAsnThrProSerSerGlnProProAlaAlaGly 368
 QY 518 GTTCTCCAGACAGAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 577
 DQ 368 sSerCysAlaTrpProThrSerAsnPheSerProLeu----- 380
 QY 578 GTGCTGCTCACCGACCTCTGCAATGCCATTGTCTGCTCTCTCTCTCTCTCTCTCT 637
 DQ 381 -----ProProSerProProLeuValSerSerProProGlyLysAla 395
 QY 638 A-----GTGCTCAATGCTGTGGAGTCACTCTCTCTCTCTCTCTCTCTCTCTCT 679
 DQ 395 aThrGlyProArgProValLeuProValCysValSerSerProValPro-----GlnMe 413
 QY 680 CCTTCAGCTCATACCGGATATGCCAGGTGCTCTCTCTCTCTCTCTCTCTCTCTCT 739
 DQ 413 tProProSerProThrAla-----ProAsnGlySerLeuAs 425
 QY 740 GAGGATGCTTTCACCAACAGCAGGTCAACAGCTCTCTCTCTCTCTCTCTCTCTCTCT 799
 DQ 425 pSerValThrTrpProValSerProProProThrSerGlyPheAlaA:apProProPr 445
 QY 800 CAGGCT 859
 DQ 445 oProProPro-----ProProProProProProProProProProProProPro 457
 QY 860 GCACCGCTCTGGCTGGCGAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904
 DQ 457 uProProProProProProProProProProProProProProProProProProPro 475
 QY 905 ACACCGACTCTGGCAGCCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 955
 DQ 476 SerProProProGlyThrProLeuAlaSerHisPheHisPheProValLeuPro 495
 QY 956 -----TCTGCT 997
 DQ 496 SerProSerAlaGlyAlaProAlaSerAlaGlnHisProLeuAsnProLeuGlyAsp 515
 QY 998 AGCCAAAGCCTACAGCATGCCCTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1057
 DQ 516 SerSerAlaSerGluProGlyLeuGlnAlaAlaSerGlnProAla-----GluSerPro 533
 QY 1058 CAGCCCGAGCTCAGCAGCTAGCTGATCATGATCATGATCATGATCATGATCATGAT 1117
 DQ 534 ThrProGln-----GlyLeuValLeuGly 541
 QY 1118 CCTGTGAGG-----CCACCGTGGGGAGCATCCAGCAGCTCTCTCTCTCTCTCTCT 1171
 DQ 542 ProProAlaProProProProProProProProProProProProProProProPro 561
 QY 1172 GAGGAGCCCCATCAACT 1231
 DQ 562 ProProProProProProPro-----ProProProProProProProProProPro 579
 QY 1232 CCTTGGAGGACATCATGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1270
 DQ 580 Pro-----ProProProProProProProProProProProProProProPro 589

RESULT 5

YAU8 SCHPO STANDARD; PRT; 354 AA.
 ID YAU8 SCHPO Q10159:
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypothetical protein C26A3.16 in chromosome I.
 GN SPAC26A3.16.
 OS Schizosaccharomyces pombe (Fission Yeast)
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Spours J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voicikert G., Aert R., Robben J., Grymonprez B.,
 RA Weetjens I., Vastreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel K., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambert R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Jesery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: TO YEAST DSK2.
 CC -1- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-aib.ch/announce/>
 CC or send an email to license@isb-aib.ch).
 CC -----
 DR EMBL; Z69240; CAA93239.1; ..
 DR PIR; T38404; T38404.
 DR GeneDB Sponbe; SPAC26A3.16; ..
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL; 2.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBO; 1.
 DR PROSITE; PS00303; UBA; 1.
 DR PROSITE; PSS0053; UBIQUITIN_2; 1.
 KW Hypothetical protein.
 FT DOMAIN 1 78 UBIQUITIN-LIKE.
 FT DOMAIN 309 353 UBA.
 FT SEQUENCE 354 AA; 36819 MW; 1A99B2D97E73A831 CRC64;

DE Mitogen-activated protein kinase 7 (EC 2.7.11.1) (Extracellular signal-regulated kinase 5) (ERK-5) (BMK kinase)

GN MAPK7 OR ERK5

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus

NCBI_TaxID=10090;

RN [1]

RA KAMAKURA S., Moriguchi T., Nishida E.

RT "Activation of the protein kinase ERK5/BMK1 by receptor tyrosine kinases: identification and characterization of a signaling pathway to the nucleus."

RL Submitted (01-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: MEK5 AND ERK5 INTERACT SPECIFICALLY WITH ONE ANOTHER AND NOT WITH MEK1/ERK1 OR MEK2/ERK2 PATHWAYS (BY SIMILARITY).

CC -!- ENZYME REGULATION: Activated by tyrosine and threonine phosphorylation (by similarity).

CC -!- DOMAIN: THE SECOND PROLINE-RICH REGION MAY INTERACT WITH ACTIN TARGETING THE KINASE TO A SPECIFIC LOCATION IN THE CELL.

CC -!- PTM: AUTOPHOSPHORYLATED ON THREONINE AND TYROSINE RESIDUES, WHEN THE C-TERMINAL PART OF THE KINASE, WHICH COULD HAVE A REGULATORY ROLE, IS ABSENT (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES. MAP KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by any for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/infocentre/> or send an email to license@ebi.ac.uk).

CC -----

DR EMBL: AB019373; BAA82039.1; --

DR HSSP: P24941; IHCL.

DR MGCD: MG11346347; Mapk7.

DR GO: GO:0045765; P:regulation of angiogenesis; (WP.

DR InterPro: IPR003527; MAP_kin.

DR InterPro: IPR00719; Prot_kinase.

DR InterPro: IPR02290; Ser_thr_pkinase.

DR Pfam: PF00369; Pkinase; I.

DR ProDom: PD03000; Prot_kinase; 1.

DR SMART: SM00220; S_TKc; 1.

DR PROSITE: PS0135; MAPK; 1.

DR PROSITE: PS0107; PROTEIN KINASE ACTIV.

DR PROSITE: PS0108; PROTEIN KINASE ACTIV.

DR PROSITE: PS50011; PROTEIN KINASE (K).

DR Transferrase: Serine/threonine-protein kinase; (cell cycle).

DR Phosphorylation.

FT DOMAIN 55 347 PROTEIN KINASE.

FT PRO-RICH 1.

FT DOMAIN 434 465 PCLY-ARG

FT DOMAIN 521 524 PRO-RICH 2.

FT NP_BIND 61 69 ATP (BY SIMILARITY).

FT BINDING 84 84 ATP (BY SIMILARITY).

FT ACT_SITE 182 219 BY SIMILARITY.

FT MOD_RES 219 219 PHOSPHORYLATION (ACTIVATES THE KINASE).

FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE).

FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE).

FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE).

FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE).

FT MOD_RES 221 221 PHOSPHORYLATION (ACTIVATES THE KINASE).

71 GTGAAGCTGGCTGACCAAGCCACTTACTTCCAAAGTCTATTCTTGGTGGCAGACAGACAA 130

449 ValAspLeuThrLeuGlnProAlaProAlaSerGluLeuAlaProProLeuArgGlu 468

131 -----CTGGCAGAAATACCTCCAGGGGCTATAGTATTTCATTTCTTGAAGCAGCTTATT 184

469 GlyAlaIleSerAspAsnThrLysAlaAlaLeuLysAlaAlaLeuLeuLysSerLeuArg 488

185 OCTGGCAAACTCCAGGAGTCTGTTCCAGACCCCTGAGCTGATTGATCTGATCTACTGTGGT 244

489 Ser

245 CGGAAGCTAAAGATGACCAACACACTTGACTTCTATGGCATTCAACCTGGTCCACTGTC 304

490 ---ArgLeuArgAspGlyProSer

305 CATGTTCTGCGAAAGTCTCGCTGCTGACCTGATCAGAAACCGGAACCTGTGGACAAAGTG 364

497 -----AlaProLeuGluAlaProGluProAlaArgLysProVal 508

365 GCTGCC-----ATGAGAGAGTTCGGGTGTTCACACTGCCCTGCACAGCAGCTCC 415

509 ThrAlaGlnGluArgGlnArgGluArgGluLysArgArgArgGlnGluArgAla 528

416 TCTTACACGGAGGGGCTTTAAGATGCTCAGCAATAAGAGTCTCTGGATCAGATCATTT 475

529 LysGluArgGlu-----LysArgGlnGluArgGluArgLysGluArgGlyAla 545

476 GTGGCCACCCAGGCTCAGCAGTACCTTCTCTTGGGGTCTTCCAGGACAGAC 535

546 GlyThrLeuGlyGlyProSerThrAspProLeuAlaGlyLeuValLeuSerAspAsnAsp 565

536 CTCTTCTCTGCTCTCGGTGATCCCAATATGCTTATGCTTGTGCTGCTGACCCAGCC 595

566 -ArgSerLeu-----LeuGluArgTrpThrArgMetAlaArgPr 578

596 CTGGTCAATGCCATTGCTCTGGTCTGCACCTCCCTAGCAGCAGTCCCAATCCCTGGG 655

578 O-----ProAlaProAlaProAlaProAlaProAlaProAla 588

656 ACTGACTCTCTCTCCGAGCATCCCTCCAGCTCATACCCGGGATATGCCAGTGGCTTC 715

588 aProAlaProAlaProSerSerAlaGlnProThrSerThrProThrGlyProVal----- 606

716 CTGTTTGAAGGGCTCTCAGATGATGAGGATGACTTCCACCAACACAGCAGTCCA---CA 772

607 -----SerGlnSerThrGlyProLeuGlu 614

773 CCTCTAAGAGTACTCCAGCTCCGCGCCAG-----CCTCCCTGGGGTACAGT 820

614 nProAlaGlySerIleProGlyProAlaSerGlnProValCysProPro----- 630

821 GGAGCTGTGGGCGCGCGCCCATCCACAGAGTGGCCAGCGCTTGGCCCTGGCCGCC 880

631 -----ProGlyProValProGlnProAlaGlyProIleProAlaPro----- 644

881 AGCACTCCGAGAGAGCTCTCACACCCAGTCTCTGGCACCAGGCTCATTTCTCTCAGGG 940

645 ---LeuGlnThrAlaProSerThrSerLeuLeuAlaSerGlnSerLeuValPro----- 661

941 ACCTCACCAATGCTCTGTTGTTCCAGTCCAGGAGCGCCATCCCAATGATCTCTTCAGC 1000

662 -----ProSerGlyLeuProGlySerGlu 669

1001 CAAGCCCTACAGCATGCGCTTTCAGGCTCTGGCAGCGCCAGCTTTCAGAGCAGTGGCAG 1060

669 yAlaProGluValLeuProTyPheProSerGlyProPro-ProProAspProGlyLeuT 689

1061 CCCAGCTGCAGCAGCTAGCTGACATCCGACATCCAGGACGATGAGCTGAGCTCGGGCCC 1120

689 hrProGlnProSerThrSer-Glu----- 696

1121 TGCAGGCCACCGTGGGGACATCCAGCAGCGCTGGAGCTCATCTTTCTGGAGGAGGCC 1180

Alignment Scores:

Pred. No.: 0.0652 Length: 806
Score: 145.00 Matches: 92
Percent Similarity: 32.41% Conservative: 37
Best Local Similarity: 23.12% Mismatches: 144
Query Match: 6.30% Indels: 125
DB: 1 Gaps: 14

US-09-735-251-3 (1-1273) x MK07_MOUSE (1-806)

2275	Db	SerSerAspLeuSerSerTrpGlyThrSerLeuLeuAlaLeuAspProAsnAlaIleGlu	2299	Db	LeuAlaLeuAspProAsnAlaIleGlu
359	QY	AAAGTGCCTGCCATGACAGAGTTCCGGGTGTGCACACTGCCTGCACAGCAGCTCTCT	418	QY	AAAGTGCCTGCCATGACAGAGTTCCGGGTGTGCACACTGCCTGCACAGCAGCTCTCT
2295	Db	AsnAlaCysLeuThrThrGlnLeuGluIleLeu--SerGlyLeuValAlaSerLysLeu	2313	Db	AsnAlaCysLeuThrThrGlnLeuGluIleLeu--SerGlyLeuValAlaSerLysLeu
419	QY	TACAGGAGGCG--	442	QY	TACAGGAGGCG--
2314	Db	LeuAlaProAlaProCysLeuIleValLeuAspProSerMetArgValIleLysVal	2333	Db	LeuAlaProAlaProCysLeuIleValLeuAspProSerMetArgValIleLysVal
443	QY	CTCAGCAATAGGAGCTCT--	472	QY	CTCAGCAATAGGAGCTCT--
2334	Db	LeuTrpGluSerGluSerProProAsnAspLeuValIleThrLeuAlaGluAspGluIle	2353	Db	LeuTrpGluSerGluSerProProAsnAspLeuValIleThrLeuAlaGluAspGluIle
473	QY	ATTGTGCCACCCAGCCTCTACAGCT--	499	QY	ATTGTGCCACCCAGCCTCTACAGCT--
2354	Db	IleAlaGluLeuProTyrLeuAsnAlaAspAspLeuLeuProMetAsnProAsp	2373	Db	IleAlaGluLeuProTyrLeuAsnAlaAspAspLeuLeuProMetAsnProAsp
500	QY	GACCTATTGCTCTGGGGTTCTCCAGGCAAGGACCTC--	541	QY	GACCTATTGCTCTGGGGTTCTCCAGGCAAGGACCTC--
2374	Db	AspProIleTyrThrArgValIleSerGlyThrAsnIleProThrAlaThrThrGluGly	2393	Db	AspProIleTyrThrArgValIleSerGlyThrAsnIleProThrAlaThrThrGluGly
542	QY	TCGTCTTCGCTGATCCCAATATGCTTGATCATCTGCTGCTCTCACCGCCCTGCTC	601	QY	TCGTCTTCGCTGATCCCAATATGCTTGATCATCTGCTGCTCTCACCGCCCTGCTC
2394	Db	SerLeuPheAlaAsp--GlnGlnLeuGluPheLeuArgProGluSer	2408	Db	SerLeuPheAlaAsp--GlnGlnLeuGluPheLeuArgProGluSer
602	QY	AATGCCATTGTCTGCTCTGCACTCCGTAGCAGGCGATGCCCAATGCTGGGACTGAC	661	QY	AATGCCATTGTCTGCTCTGCACTCCGTAGCAGGCGATGCCCAATGCTGGGACTGAC
2409	Db	AsnProPheProPheAlaSerHisAspSerSerGlnSerLeuAspValProSerSerPro	2428	Db	AsnProPheProPheAlaSerHisAspSerSerGlnSerLeuAspValProSerSerPro
662	QY	TCCTCTTCGGGAGCATGCCCTCCACCTCATACCGGGATATGCCAGGTGGCTTCCTGTTT	721	QY	TCCTCTTCGGGAGCATGCCCTCCACCTCATACCGGGATATGCCAGGTGGCTTCCTGTTT
2429	Db	SerSerGly--SerAspLysTyrGluGluAspProThrGlyIleValTyr	2444	Db	SerSerGly--SerAspLysTyrGluGluAspProThrGlyIleValTyr
722	QY	GAAGGGCTCTCAGATGAT--	739	QY	GAAGGGCTCTCAGATGAT--
2445	Db	AspAlaProValAspAspMetSerAspMetAlaMetAsnLysAlaLysAlaTrpGlnGlu	2464	Db	AspAlaProValAspAspMetSerAspMetAlaMetAsnLysAlaLysAlaTrpGlnGlu
740	QY	-----GAGGATGACTTTCACCCAAACACACAGGTCACACCCCTCT	778	QY	-----GAGGATGACTTTCACCCAAACACACAGGTCACACCCCTCT
2465	Db	TrpLeuGluAspGlyPheAlaGluAspAspTyrArgGluLeuSerAsnAlaMetPro--	2483	Db	TrpLeuGluAspGlyPheAlaGluAspAspTyrArgGluLeuSerAsnAlaMetPro--
779	QY	AGCAATCTCCACGCTCCGCCCCAGCCTCCCTGGGTACAGTGGAGCTGGCGCCCGG	838	QY	AGCAATCTCCACGCTCCGCCCCAGCCTCCCTGGGTACAGTGGAGCTGGCGCCCGG
2484	Db	-----AlaProProLys	2487	Db	-----AlaProProLys
839	QY	-----CCCATCACCACAGATGAGCTGGCCACGCCCTTGCCCTGGCCACACCTCCGAG	892	QY	-----CCCATCACCACAGATGAGCTGGCCACGCCCTTGCCCTGGCCACACCTCCGAG
2498	Db	ThrThrProValGluSerLys--	2499	Db	ThrThrProValGluSerLys--
893	QY	AGCAGCTCTCACACACGACTCTCTGGACCCACGGGTATCTCTCAGGAGCTCACCAATG	952	QY	AGCAGCTCTCACACACGACTCTCTGGACCCACGGGTATCTCTCAGGAGCTCACCAATG
2500	Db	SerValAspArgAlaProThrLeuProProLysAlaAlaProLeuProProSerAspAl	2519	Db	SerValAspArgAlaProThrLeuProProLysAlaAlaProLeuProProSerAspAl
953	QY	TCCTCTGGTCTCCAGTCAGGACGCCCA--	1000	QY	TCCTCTGGTCTCCAGTCAGGACGCCCA--
2519	Db	AsnAlaIleMetSer--GlyLysProValPheLysTyrThrProGlyAsnLysSerAl	2538	Db	AsnAlaIleMetSer--GlyLysProValPheLysTyrThrProGlyAsnLysSerAl
1001	QY	CAAGCCCTACAGCATGCCCTTTCAGGCTCTGGGACGCCACCGCTTCAGAGCCAGTGGCAG	1060	QY	CAAGCCCTACAGCATGCCCTTTCAGGCTCTGGGACGCCACCGCTTCAGAGCCAGTGGCAG
2538	Db	aValProProSerValPro--	2553	Db	aValProProSerValPro--
1061	QY	CCCCAGCTCAGCAGCTACGTGACATGGGATCCAGGACGATGAGCTGAGCTGGCGGCC	1120	QY	CCCCAGCTCAGCAGCTACGTGACATGGGATCCAGGACGATGAGCTGAGCTGGCGGCC
2553	Db	roProLeuProGlnSerThrSerLysAlaAlaSerGly--	2568	Db	roProLeuProGlnSerThrSerLysAlaAlaSerGly--
1121	QY	TGCAGGCAACCGGTGGGACATCCAGCAGCCCTGGAGCTTCATCTTTGCTGGAGGAGCCC	1180	QY	TGCAGGCAACCGGTGGGACATCCAGCAGCCCTGGAGCTTCATCTTTGCTGGAGGAGCCC
2569	Db	ThrLeuProProAlaProProLeuProGlnSerThrSerLysAlaAlaSerGlyProPro	2588	Db	ThrLeuProProAlaProProLeuProGlnSerThrSerLysAlaAlaSerGlyProPro
1181	QY	CATGAATCTCCTGCTTCCCTCGAACCCCAAGTTCGACAGGCTACTGCCCTTGGGAG	1240	QY	CATGAATCTCCTGCTTCCCTCGAACCCCAAGTTCGACAGGCTACTGCCCTTGGGAG

Db 2589 ProThrLeuProProAlaProProLeuProGlnSerThrSer-----Lys 2603

Qy 1241 GCACCTCATGAGTGCTCCATCTCTCCCT 1270

EL 2604 AlaAlaSerGlyProProThrLeuPro 2613

RESULT 9

N153_RAT STANDARD; PRT: 1468 AA.

AC P49791;

DT 01-OCT-1996 (Rel. 34, Created;

DT 01-OCT-1996 (Rel. 34, Last sequence update);

DT 28-FEB-2003 (Rel. 41, Last annotation update);

DE Nucleolar pore complex protein Nup153 [Nucleoporin Nup153] (153 kDa

DE nucleoporin).

GN NUP153.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;

OX NCBI_TaxID=10116;

RN 1.

RP SEQUENCE FROM N.A., AND SEQUENCE OF 602-613, 622-645 AND 971-993.

RC STRAIN=Buffalo; TISSUE=Liver;

RX MEDLINE=93137325; PubMed=8422679;

RA Sukegawa J., Blobel G.;

R1 "A nucleolar pore complex protein that contains zinc finger motifs,

R1 binds DNA, and faces the nucleoplasm.";

RL Cell: 72:29-38(1993).

CC -!- FUNCTION: Possible DNA-binding subunit of the nuclear pore

CC complex (NPC). The repeat-containing domain may be involved in

CC anchoring components of the pore complex to the pore membrane.

CC -!- SUBUNIT: Interacts with SENP2 [By similarity].

CC -!- SUBCELLULAR LOCATION: Nuclear pore complex. Located to the

CC terminal ring structure of the nucleoplasmic cage.

CC -!- DOMAIN: Contains F-X-F-G repeats.

CC -!- PTM: Phosphorylated in interphase, hyperphosphorylated during

CC mitosis. May play a role in the reversible disassembly of the

CC nuclear pore complex during mitosis.

CC -!- SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS

CC SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST

CC NUP1, NUP1, POM 12; AND MAMMALIAN P62.

CC -!- SIMILARITY: Contains 4 RANBP2-type zinc fingers.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed, changed, copied and for commercial

CC entities requires a license agreement (See: http://www.ebi.ac.uk/submit/submit.html)

CC Or send an email to license@ebi.ac.uk.

CC

DR EXBL: L06821; NOT ANNOTATED_CDS.

DR PIR: A44345; A44345.

DR InterPro: IPR001876; Znf_RangDP.

DR Pfam: PF01093; Nucleoporin_FG_9.

DR Pfam: PF00641; zf-RanBP; 4.

DR SMART: SM00547; Znf_RBZ; 4.

DR PROSITE: PS01358; ZF_RANBP2_1; 4.

DR PROSITE: PS01093; ZF_RANBP2_2; 4.

KW Nuclear protein; Transport; Repeat; Zinc-finger; DNA-binding;

KW phosphorylation.

KW DOMAIN 4 15 GLY-RICH.

FT DOMAIN 442 446 POLY-GLY.

FT ZN_FING 657 687 RANBP2-TYPE 1.

FT ZN_FING 721 750 RANBP2-TYPE 2.

FT ZN_FING 789 818 RANBP2-TYPE 3.

FT ZN_FING 846 875 RANBP2-TYPE 4.

SQ SEQUENCE 1468 AA; 152823 MW; C3DFD9697C556A7C CRC64;

Alignment Scores:

Pred. No.: 0.0761 Length: 1468

Score: 144.50 Matches: 106

Percent Similarity: 36.82% Conservative: 49

Best Local Similarity: 25.18% Mismatches: 144

Query Match: 6.28% Indels: 122

DB: 1 Gaps: 20

US-09-735-251-3 (1-1273) x N153_RAT (1-1468)

Qy 149 GGGGCTAT-----AGTATTTCATTCTGAAGCAGCTATT 184

Db 1073 GlyGlyPheThrPheGlyLysValAspSerAlaAlaLeuSerSerProSerMetPheVal 1092

Qy 185 GCTGGCAAACTCCAGGAGTCTGTCAGACCCCTGAGCTGATTCATCTGCTGCTGCT 244

Db 1093 LeuGlyValGThrGluGluLysGlnGluProValThrSerThrSerLeuValPheGly 1112

Qy 245 CGGAAAGCTA-----AAAGATGAC 262

Db 1113 LysLysAlaAspAsnGluGluProLysCysGlnProValPheSerPheGlyAsnSerGlu 1132

Qy 263 CAGACACTTGATCTTATGGCATTCAACTGGGTCACATGTCATGTCGGAAGTCC 322

Db 1133 GlnThrLysAspGluSerSerLysProThrPheSerPheSerValAlaLysProSer 1152

Qy 323 TGGCTTGAACCTGATCAGAAACCGGAA----- 349

Db 1153 ValLysGluSerAspGlnLeuAlaLysAlaThrPheAlaPheGlyAsnGlnThrAsnThr 1172

Qy 350 CTTGTGGCAAAAGTGGCTGCCATGAGAGAGTTCCTGGGTGTTGCACACTGCCCTGCACAGC 409

Db 1173 ThrThrAspGlnGlyAlaLysProAlaPheSerPheLeuAsnSerSerSerSer 1192

Qy 410 AGCTTC-----TCCTACAGGAGGCGGTCTTAAGATGCTCAGCAATTAAGGAG 457

Db 1193 SerSerThrProAlaThrSerSerSerAlaSerIlePheGlySer---SerThrSerSer 1211

Qy 458 TCTCTGGATCAGATCATTTGTGGCCACCCAGCCCTCAGAGTGACCCCTATTGCTCTCTGG 517

Db 1212 SerSerProProValAlaAlaPheValPheGlyGlnAlaSerAsnProValSer----- 1229

Qy 518 GTTCTCCAGGACAGGACCTCTCTCTCTGCTGATCCCAATATGCTTCATACGTTG 577

Db 1230 -----SerSerAlaPheGlyAsnSerAlaGluSerSerThrSer 1242

Qy 578 GTGCT-----GCTCACCAGCCCTCGTCAATGCCATTCCTCTGTT 619

Db 1243 GlnProLeuLeuPheProGlnAspGlyLysProAlaThrThrSer----- 1257

Qy 620 CTGCATCTGGTAGGAGGAGTCCCA-----ATGCTGGGACGACGCTCTCTCTCCCG 673

Db 1258 -----SerThrAlaSerAlaAlaProPheValPheGlyThrGlyAlaSerSerAsn 1275

Qy 674 AGCATGCTCTCCAGCTACATACCGGATATGCCAGGTGGCTCTCTGTTGAAGGCTCTCA 733

Db 1276 SerThrValSerSer-----GlyPheThrPheGlyAlaThrThr 1288

Qy 734 GATGATGAGGATGACTTTCACCCAAACACAGGTCACACCCCTCTAGC----- 781

Db 1289 -----ThrSerSerSerSerGlySerPheValPhe 1299

Qy 782 -----AGTACTCCAGCTCCCGCCAGCC-----TCC 808

Db 1300 GlyThrGlyHisSerAlaProSerAlaSerPheGlyAlaAsnGlnThrProThr 1319

Qy 809 CTGGGTACAGT---GGAGCTCTGGGCGCCCGCCATCACCCAGAGTGAGTGCCACCC 865

Db 1320 PheGlyGlnSerGlnGlyAlaSerGlnProAsnProProSerPheGlySerIleSerSer 1339

Qy 866 GCCTTGGCCCTGGCCAGC-----ACTCCG 889

Db 1340 SerThrAlaLeuPheSerAlaGlySerGlnProValProProThrPheGlyThrVal 1359

Qy 890 GAGAGCAGCTCTCACACCCAGCT-----CTGGACCCAGGCTCATTCCTCA 937

Db 1360 SerSerSerGlnProProValPheGlyGlnGlnProSerGlnSerAlaPheGlySer 1379

QY 938 GGGACCTCACCAATGCTCTGCTGCTCAGTCAGGAGCCCATCACCAGATGATCTCTC 997
 D 1380 GlyThrAlaAsnAlaSerValPheGlnPheGlySerThrThrAsn-----Phe 1397
 QY 998 AGCAAGCCCTACAGCATGCGCTTCAGGCTC-----TGAGCAGCCAGGCTTCAGAGC 1051
 D 1398 AsnPheThrAsnAsnAsn--ProSerGlyValPheThrPheGlyAlaSerProSe: 1415
 QY 1052 CAGTGGCAGCCCACTGACAGCAGTCTAGTGTGATGAGCAGTCTGAGCAGTCTGAGC 1111
 D 1416 -----ThrProAlaAlaAlaGlnProSerGlySerGlyGlyPheSerPheSer 1432
 QY 1112 CTGCGCGCTGACGAGCCAGCCGCTGGGAGATCTT AAGCAGCCTGAGTCTCATCT 1165
 D 1433 GlnSerProAlaSerPheThrValGlySerAsnGlyAsnMetPheSerSerSer: 1491
 RESULT 10
 DAB1_HUMAN STANDARD; PRT; 555 AA.
 AC 075553; Q9NVAB.
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Disabled homolog 1.
 GN DAB1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99009343; PubMed=9790777;
 RA Lambert de Rouvroit C., Goffinet A.M.
 RT "Cloning of human DAB1 and mapping to chromosome 1p32 p32";
 RL Genomics 53:246-247(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Fazili Z., Sun X., Xu X.-X.;
 RT "Aberrant disabled-1 expression in tumors";
 RL Submitted (MAY-2000) to the EMBL/GenBank/CDR databases.
 CC -!- FUNCTION: Adapter molecule functioning in neural development. (By
 similarity).
 CC -!- SUBUNIT: Associates with the SH2 domains of Src, Fyn and Abl (by
 similarity).
 CC -!- DOMAIN: THE PID DOMAIN STRUCTURALLY RESEMBLES TO THE ASN PRO-XAA
 TYR(P) MOTIF FOUND IN MANY TYROSINE PHOSPHORYLATED PROTEINS.
 CC -!- PTM: Phosphorylated on Tyr 148 and 149 and upon insulin induction,
 in embryonic neurons (By similarity). Also phosphorylated on Ser-
 491 independently of insulin stimulation (By similarity).
 CC -!- SIMILARITY: Contains 1 PID domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF071062; AAC70068.1;
 CC EMBL; AF263547; AAF73058.1;
 CC Genew; HGNC:2661; DAB1.
 CC MIX; 603448;
 CC InterPro; IPR006020; PTB_P1D.
 CC Pfam; PF00640; PID; 1.
 CC SMART; SM00462; PTB; 1.
 CC PROSITE; PS01179; PID; 1.
 CC Developmental protein; Neurogenesis; Phosphorylation.
 KW DOMAIN 36 189 PID.
 FT MOD_RES 198 198 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 220 220 PHOSPHORYLATION (BY SIMILARITY).
 FT MOD_RES 491 491 PHOSPHORYLATION (BY CDK5) (BY

FT CONFLICT 86 86
 FT CONFLICT 240 241
 FT CONFLICT 291 291
 FT CONFLICT 298 298
 FT CONFLICT 326 326
 FT CONFLICT 335 335
 FT CONFLICT 379 379
 FT CONFLICT 553 553
 SC SEQUENCE 555 AA; 59979 MW; A4B133A0C3AF2034 CRC64;
 Alignment Scores:
 Pred. No.: 0.077 Length: 555
 Score: 143.50 Matches: 87
 Percent Similarity: 35.94% Conservative: 51
 Best Local Similarity: 22.66% Mismatches: 143
 Query Match: 6.23% Indels: 103
 DB: 19
 US-09-735-251-3 (1 1273) x DAB1_HUMAN (1-555)
 QY 242 GGTGGGAGTAAGATGACGACGACATCTGATCTTATGGCATTCACTGGGTCCACT 301
 D 57 GlyAspLysLeuCysGlnAspSerMetMetLysLeuLysGlyValValAlaGlyAla--- 75
 QY 302 GTCCATCTTCTGCGAAAGTCTCGCTGACCTGATCAGAAACCCGAAACCTGTG----- 355
 D 76 -----ArgSerLysGlyGluHisGlyGlnLysLeuPheLeuThrLeuSer 90
 QY 356 -----GACAAAGTGGCTGCCATGAGAGAGATTCGGGTGTTG 391
 D 91 PheGlyGlyLeuLysLeuPheAspGluLysThrGlyAlaLeuGlnHis----- 106
 QY 392 CACACTGCGCTGACAGCAGCTCTCTTACAGGAGGCGGTC-----TTT 436
 D 107 HisAlaValHisGluLeuSerTyrLeuAlaLysAspLeuThrAspHisArgAlaPhe 126
 QY 437 AAGATGTCAGCAATAAGGAGTCTCTGATCAGATCATCTGTGGCACCACCCAGCCCTCAGC 496
 D 127 GlyTyrValCysGlyLysGluGlyAsnHisArgPheValAlaLeuLysThrAlaGlnAla 146
 QY 497 AGTGACCTATTGCTCTTGGGGTCTCAGGACAGGACCTCTTCTGCTTGGCTGAT 556
 D 147 AlaGluProVal-----IleLeuAspLeuArgAspLeuPheGlnLeuLysTyrGlu 163
 QY 557 CCGAATATGTTGATGCTG-----GTGCTGCTCACCAGCC 595
 D 164 LeuLysHisArgGluGluLeuGluLysAlaGlnLysAspLysGlnCysGluGlnAla 183
 QY 596 CTCCTCAATGCCATTGTC-----CTGGTTCG 622
 D 184 ValTyrGlnThrIleLeuGluGluAspValGluAspProValTyrGlnTyrIleValPhe 203
 QY 623 CACTCCCTAGCAGCAGTCCCAATG-----CTGGGACTGACTCTCTTCCCGGAGC 676
 D 204 Glu---AlaGlyHisGluProIleArgAspProGluThrGluGluAsnLeuTyrGln 221
 QY 677 ATGCGCTCCAGCTCATACCGGATATCCAGTGGCTTCTCTGTTTGAAGGCTCTCAGAT 736
 D 222 ValProThrSerGlnLysLys-----GluGlyValTyrAsp 233
 QY 737 GATCAGGATGACTTTCACCCAAACACCGATCCACCTCTAGCAGTACTCCAGCTCC 796
 D 234 -----ValProLysSer 237
 QY 797 CGCCAGCCTCTCCCTGGGGTACAGTGGAGTCTGTGGGCCCGGCCCATCACCCAGAGTGA 856
 D 238 GlnProValSer-----AlaValThrGlnLeuGlu 247
 QY 857 CTGCGCCACCGCTTGGCCCTGGCCGACGACTCTCCGAGGAGGAGCTCTCACACCGACTCT 916
 D 248 LeuPheGlyAspMet-----SerThrProAspLeuThrSerProThrPro 264

QY 917 GGCACCCAGGT-----CATCTCAGGACCTCACCACATGTCTCTGTGT 961
 Db 265 AlatrProGlyAspAlaPheIleProSerSerGlnThrLeuProAaSerAlaAsp 284
 QY 962 GTCCAGTCAGGAGCCCATCACCACATGATCTCTCAGGCAAGCCCTACAGATGCCCT 1021
 Db 285 ValPheSerValPro-----LeuGlyThrAlaAaValProGlyTyr 300
 QY 1022 CAGCCCTCTGGCAG-----CCAGCCCTTCAAGCCAGTGGCAGCCCGAGCTGCACAG 1075
 Db 301 ValAlaMetGlyAaValLeuProSerPheThrGlyGln---GlnProLeuValGlnGln 319
 QY 1076 CTAGCTGACATGCGCATCCAGSAGATGAGCTGAGCTGGCGCTCCAGCCACCCGCTG 1135
 Db 320 GlnMetValMetGlyAaHisProProValAlaGlnValMetProGlyAaGlnProIle 339
 QY 1136 GGCACATCAAGCAGCCCTGCGAGCTCATCTTCTGCGAGGAGCCCATGAACCTCCCTGT 1195
 Db 340 AlatrProGly---GlnProGlyLeuPheProAlaThrGlnGlnProThrValAla 358
 QY 1196 TCCCTGTAACCCCGCAGCAAGTGGAGAGGCAATGCTTGGAGGAGCCTCATGAAGGTG 1255
 Db 359 GlyGlnPheProProAaAala-----PheValProThrGlnThrValMetProLeu 375
 QY 1256 CCTCCATCTCTC 1267
 Db 376 ProAlaAaMet 379

RESULT 11
 SREC_HUMAN
 ID_SREC_HUMAN STANDARD: FRT: 830 AA
 AC Q14162.043701:
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 29-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Endothelial cells scavenger receptor precursor (Acetyl LDL receptor;
 DE Scavenger receptor class F member 1).
 GN SCARF1 OR SREC OR KIAA149.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN 11_TaxID=9606;
 RP SEQUENCE FROM N.A.
 RC Tissue: Umbilical vein endothelial cells.
 RX MEDLINE=98058697; PubMed=9395444;
 RA Adachi H., Tsujimoto M., Arai H., Imase K.
 RT "Expression cloning of a novel scavenger receptor from human
 RT endothelial cells".
 RL J. Biol. Chem. 272:31217-31220(1997).
 RN [2].
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22086180; PubMed=11978792;
 RA Adachi H., Tsujimoto M.;
 RT "Characterization of the human gene encoding the scavenger receptor
 RT expressed by endothelial cell and its regulation by a novel
 RT transcription factor, endothelial zinc finger protein-2".
 RL J. Biol. Chem. 277:24014-24021(2002).
 RN [3].
 RP SEQUENCE FROM N.A.
 RC Tissue: Bone marrow;
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nozura N.;
 RT "Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (K1A0121-K1A0160) deduced by
 RT analysis of cDNA clones from human cell line KG-1".
 RL DNA Res. 2:167-174(1995).
 CC -!- FUNCTION: Mediates the binding and degradation of acetylated low
 CC density lipoprotein (Ac-LDL). Mediates heterophilic interactions,
 CC suggesting a function as adhesion protein (By similarity).
 CC -!- SUBUNIT: Heterophilic interaction with SREC2 via its extracellular
 CC domain. The heterophilic interaction is suppressed by the presence
 CC of ligand such as Ac-LDL (By similarity).

CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
 CC -!- TISSUE SPECIFICITY: Endothelial cells.
 CC -!- SIMILARITY: Contains 6 EGF-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: D86864; BA24070.1;
 DR EMBL: AB052946; BAC02692.1;
 DR EMBL: D63483; BAC09770.1;
 DR HSSP: P01180; 2BN2.
 DR Genew: HGNC:16820; SCARF1.
 DR GO: GO:0016021; C:integral to membrane; IDA.
 DR GO: GO:0030169; F:low-density lipoprotein binding activity; IDA.
 DR GO: GO:0004888; F:transmembrane receptor activity; TAS.
 DR GO: GO:0045192; P:low-density lipoprotein catabolism; TAS.
 DR GO: GO:0006899; P:receptor mediated endocytosis; TAS.
 DR InterPro: IPR006209; EGF-like.
 DR InterPro: IPR006210; IEGF.
 DR InterPro: IPR020049; Laminin_EGF.
 DR PRINTS: PR00011; EGF_LAMININ.
 DR SMART: SM00181; EGF; 5.
 DR PROSITE: PS00022; EGF_1; 6.
 DR PROSITE: PS01186; EGF_2; 6.
 DR Cell adhesion: Receptor; Repeat; Signal; Transmembrane;
 KW EGF-like domain; Glycoprotein; POTENTIAL.
 FT SIGNAL 1..59
 FT CHAIN 20..830
 FT DOMAIN 20..421
 FT TRANSMEM 422..442
 FT DOMAIN 443..830
 FT DOMAIN 53..87
 FT DOMAIN 95..130
 FT DOMAIN 155..191
 FT DOMAIN 215..249
 FT DOMAIN 302..339
 FT DOMAIN 351..382
 FT DOMAIN 476..520
 FT DOMAIN 622..798
 FT DOMAIN 830..830
 FT DISULFID 57..69
 FT DISULFID 63..75
 FT DISULFID 77..86
 FT DISULFID 99..111
 FT DISULFID 105..118
 FT DISULFID 120..129
 FT DISULFID 159..172
 FT DISULFID 165..179
 FT DISULFID 181..190
 FT DISULFID 219..230
 FT DISULFID 225..237
 FT DISULFID 239..248
 FT DISULFID 306..319
 FT DISULFID 323..326
 FT DISULFID 329..338
 FT DISULFID 355..363
 FT DISULFID 358..370
 FT DISULFID 372..381
 FT CARBOHYD 382..389
 FT CARBOHYD 389..392
 FT CARBOHYD 393..393
 FT CONFLICT 662..662
 SQ SEQUENCE 830 AA; 87430 MW; F560D9E1A64D779 CRC64;
 Alignment Scores:
 Pred. No.: 3.0875 Length: 830
 Score: 143.00 Matches: 128
 Percent Similarity: 26.79% Conservative: 40

NCBI TaxID=9606:

SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
MEDLINE#20021762; PubMed#13552926;
Liarg Y., Wang A., Belvanteva I.A., Anderson D.W., Probst F.J.,
Barber T.D., Miller W., Touchan J.W., Jin L., Sullivan S.L.,
Sellers J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman I.B.,
Fridell R.A.;
"Characterization of the human and mouse unconventional myosin XV
genes responsible for hereditary deafness DFNB3 and shaker 2,"
Genomics 61:243-258(1999).

```

EMBL: AF144095; AAC05903.1;
EMBL: AF051975; -; NOT_ANNOTATED;
PIR: A59266; A59266.
HSSP: P10587; 1BR2.
GeneW: HGNC:7594; MYC15A.
MIM: 602666; -
YIM: 600336; -
GO: GO:0007605; P:hearing; TAS.
InterPro: IPR000399; Band 4.1.
InterPro: IPR000048; IQ region.
InterPro: IPR001609; myosin_head.
InterPro: IPR000857; MYTH4.
InterPro: IPR001452; SH3.
Pfam: PF00612; IQ_3.
Pfam: PF00663; myosin_head; 1.
Pfam: PF0784; MYTH4; 2.
PRINTS: PR00193; MYOSINHEAVY.
ProDom: PD000355; myosin_head; 1.
SMART: SM00295; B41; 1.
SMART: SM00015; IQ; 3.
SMART: SM00242; MYSC; 1.
SMART: SM00139; MYTH4; 2.
SMART: SM00326; SH3; 1.
PROSITE: PS00560; FERM_1; FALSE_NEG.
PROSITE: PS00661; FERM_2; FALSE_NEG.
PROSITE: PS20057; FERM_3; 1.

```

DR PROSITE; PS50096; IQ; 3.
 DR PROSITE; PS500C2; SH3; FALSE NEG.
 KW MYOSIN; ATP-binding; Actin-binding; Coiled coil; Repeat; SH3 domain;
 KW Calmodulin-binding; Disease mutation; Deafness.
 FT DOMAIN 1 1887 HEAD OR MOTOR DOMAIN
 FT DOMAIN 1888 2329 NECK OR REGULATORY DOMAIN.
 FT DOMAIN 2033 3330 TAIL.
 FT DOMAIN 3233 3350 COILED COIL (POTENTIAL).
 FT DOMAIN 1792 1799 ACTIN-BINDING (POTENTIAL).
 FT DOMAIN 1902 1924 IQ 1.
 FT DOMAIN 1925 1954 IQ 2.
 FT DOMAIN 1955 1976 IQ 3.
 FT DOMAIN 2867 2953 SH3.
 FT DOMAIN 3209 3530 FERM.
 FT NP BIND 1315 1322 ATP (POTENTIAL).
 FT VARIANT 2111 2111 N -> Y (IN DFNB3; FAMILY FROM BENGALA).
 FT VARIANT 2113 2113 /FTID=VAR_010303;
 FT VARIANT 2113 2113 N -> Y (IN DFNB3; FAMILY FROM BENGALA).
 FT VARIANT 2113 2113 /FTID=VAR_010304;
 FT VARIANT 2113 2113 I -> F (IN DFNB3; INDIAN FAMILY).
 SQ SEQUENCE 3530 AA; 395171 MW; 3D103923D4BCBE4A CRC64;
 Alignment Scores:
 Pred. No.: 0.115 Length: 3530
 Score: 142.50 Matches: 95
 Percent Similarity: 36.31% Conservative: 35
 Best Local Similarity: 26.54% Mismatches: 138
 Query Match: 6.19% Indels: 90
 DB: 1 Gaps: 19
 JS-09-735-251-3 (1-1213) x MY15_HUMAN (1-3530)
 QY 296 TCACCTGTCATGTTCTGCGAAGTCTCTGGCCCTGAACCTGATCAGAAACCGGAACCTGTG 355
 DB 571 SerLeuA:aa:gPhLeuLysLysThrLeuSerGlu.....LysLysProIle 586
 QY 356 GACAAGTGGTGGCATGAGAGAGTTCGGGTGTTCACACTGCCCTGCACAGCAGCTCC 415
 DB 587 AlaA:gLeuA:gGlySerGlnLysThrArgAla.....GlyGlyPro 600
 QY 416 TCTTACAGGGAGCGGTCTTAAGATGCTCAGCAATGAGAGTCTCTGGATCAGATCATTT 475
 DB 601 AlaValA:gGluAlaAlaA:gLysArgPheGlyTyrLys.....LeuAla 615
 QY 476 GTGGCAACCCAGGCGCTCAGCAGTGAACCTATTGCTCTGGGGTCTCCAGGACAAGGAC 535
 DB 616 GlyMet:AspProG:LysProGlyThr:ProIleValLeuA:gArgAlaGln-ProArgAl 635
 QY 516 CTTCTCTCTGCTCTGGCTGCTG.....ATCCCAATATGCTTGAT 571
 DB 635 aArgSerSer:AsnAspAlaA:gArgProProAlaProGlnProAlaProArgThrLeuSe 655
 QY 572 ACCTTGGTGGCTGTCTACCCAGCGCTCGTCAATGCCATTGCTCTGGTCTGCACTCCGTA 631
 DB 655 RhIstrpSerAlaLeuLeuSerProProValProProArg..... 668
 QY 632 GCAGGCAGTGGCCCAATGCCCTGGGACTGACTCTCTTCCCGGAGCATGCCCTCCAGCTCA 691
 DB 669ProProSerSerGlyProProProAlaPro---ProLeuSerProAlaIle 684
 QY 692 TACCGGATATGCCAGGTGGCTCTCTGTTGAAGGGCTCTCAGATGATGAGATGACTTT 751
 DB 684 uSerGlyLeuProArgProAlaSerProTyr---GlySer-----LeuArgAr 699
 QY 752 CACCCAAACACAGGTCCACACCTCTAGCAGTACTCCGAGTCTCCGCGCCAGGCTCCCTG 811
 DB 699 GhIspProProTtpAlaAlaProAlaHisValProProAlaProGlnAlaSerGlyTr 719
 QY 812 GGTACAGTGGAGCTGTGGGCGCGGCCCATCACCCAGAGTCAGCTGGCCACCG---- 866
 DB 719 palApheA:gGlu-----ProProAlaValSerProGluVal-----ProProAspLe 735
 QY 867 -----CCTTGGCCCTGGCCGAGCACTCCGAGAGCAGCTCTCACACACCGCACTCT 916


```
Db      394  GluAspAspGlyGluAspLeuAspThr.....ArgProSerAspLeuMet 409
QY      605  GCC-----ATTGCTGCTGTTCTCTCAGCTCCAGCAGGAGCTGCCCAATG 649
Db      410  AlaGlnMetCysProGlnLeuValAlaThrAlaPheSerLysAspGlySerAlaThrGlu 429
QY      650  CTTGGGACTGACTCTCTCTCCGAGCATGCTCTCAGCTCATACGGGATATGCCAGGT 709
Db      430  AspGluAspSerAsnSerAspAspThrAspGluSerThrAspAspGluAspGluAspSer 449
QY      710  GGCTTCCTGTTTGAGGGCTCTCAGATCATATGATGAC..... 748
Db      450  AspSerGluAsnAspSerAspSerGluAspGluGluGluGluGluGluGluGlu 469
QY      748  ----- 748
Db      470  ProValLysIleAlaGlyAlaLysArgGlyArgAsnAspAspGluGluValSerProLeu 489
QY      749  -----TTTACCCAAACACACAGGTCACACATCTCTAGCAGTA----- 785
Db      490  AlaLysLysProArgLlePheSerProProValArgProArgProGlnAlaIleArgVa 509
QY      786  -----CTCCAGCTCCCGCCAGGCTCCCTGGGTACAGTGGAGCTCTGGGCCCGGCC 841
Db      509  LysLeuProSerProAlaProSerSerArgGlySerThrProValSerProValSerPr 529
QY      842  ATCACCAGAGTACGCTGGCCACGGCT.....GGCCCCGGCCAGCACT 886
Db      529  SerProLysAlaLysArgProAlaGlnAlaThrSerLeuLeuAlaSerHisProMetLy 549
QY      887  CCGGAGAGCAGCTCTACACACCGACTC.....CTGGCCCGCAGGCTCATTC 934
Db      549  LysArgGluLysLeuGlnGluGluLeuArgLysAlaGlyLeuAlaProProSerAlaPr 569
QY      935  T-----CAGGGACCTCACCAATGCTCTGTTGTCAGTGGAGCGCCCATCACCAAT 988
Db      569  ValLeuMetGlyProAspGlyValProLeuGlyThrValArgSerArgSerProSe 589
QY      989  GATCTCTTCAGCCAGCCCTACAGC-----ATGCCCTTCAGGCTCTGGCGAG 1036
Db      589  rValSerSerProProSerValSerValSerLeuProLeuProSerArgGlyValProSe 609
QY      1037  CCCAGCCTTCAGAGCCAGTGGAGCCCGACCTTACACAGTACGTACATGG 1088
Db      609  rGlyGlyLleLysValThrGlyAspPro...ThrProIleValAsnArgP 624
```

Search completed: November 12, 2003 06:47:11
Job time : 41 secs

1 19:4 83.1 380 4 Q96IC3 Q96IC3 homo sapien

TISSUE=Muscle;
RA Strausberg R;
FL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC007913; AAH07913.1; 1
DR EMBL; BC030655; AAH30055.1; 1
DR EMBL; BC033919; AAH33919.1; 1
DR InterPro; IPR001901; SECE.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR PROSITE; PS01367; SECE_SEC61G; 1.
DR PROSITE; PS00553; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 380 AA; 40510 MW; 0B64DF79597E0931 CRC64;
Alignment Scores:
Pred. No.: 2,83e-139 Length: 380
Score: 1914.00 Matches: 379
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 1
Query Match: 83.15% Indels: 2
DB: 4 Gaps: 0
US-09-735-251-3 (1-1273) x Q96S82 (1-380).
QY 44 ATGCTCTCTCAGACTGGCAGCTGGCGTGAATCGGTGACGACCCACTTACTCCAAAG 103
Db 1 MetSerLeuSerAspThrHisLeuAlaValLeuAlaAspGlnProLeuThrProlys 20
QY 104 TCTATCTCTCGGTGGCCAGACAGAACTGGGAGATATCTCGTACGGGCTATATATT 163
Db 21 SerIleLeuArgLeuProGlnThrGluLeuGlyGlyTyrSerLeuGlyGlyTyrSerIle 40
QY 164 TCATTCTGAAGCAGTTTATGCTGGAACCTGACGAGTCTGTTCCAGACCCCTGAGTGT 223
Db 41 SerPheLeuGlnLeuIleAlaGlyLysLeuGlnGluSerValProAspProGluLeu 60
QY 224 ATTCATCTGATCTGCTGGAGCTTAAAGATGACACACACTTGACTTCTATGGC 283
Db 61 IleAspLeuIleTyrCysGlyArgLysLeuLysLeuAspGlnThrLeuAspPheTyrGly 80
QY 284 ATTCACCTGGGTCCACTGCTGCTGCTGGAAGTCTGCTGCTGCTGCTGCTGCTGCTG 343
Db 81 IleGlnProGlySerThrValHValHValHValHValHValHValHValHValHValH 100
QY 344 CGGACCTGTGAACAAAGT 403
Db 101 ProGluProValAspLysValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 120
QY 404 CACAGCAGCTCTCTTACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 463
Db 121 HisSerSerSerSerTyrArgGluAlaValPheLysValLeuSerAsnLysGluSerLeu 140
QY 464 GATCAGATCATTTGTGCCCACCCAGCCCTGACAGTGAACCTATTGCTCTTGGGGTTCTC 523
Db 141 AspGlnIleIleValAlaThrProGlyLeuSerSerAspProIleAlaLeuGlyValLeu 160
QY 524 CAGGACAGACCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583
Db 161 GlnAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspThrLeuValPro 180
QY 584 GTCACCCAGCCCTGCTGATGCGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 643
Db 181 AlaHisProAlaLeuValAsnAlaIleValLeuValLeuHisSerValAlaGlySerAla 200
QY 644 CCAATGCTGGGACTGACTCTCTTCCCGAGGACATGCCCTCCAGCTCATACCGGATATG 703
Db 201 ProMetProGlyThrAspSerSerArgSerMetProSerSerSerTyrArgAspXet 220
QY 704 CCAGGTGGCTTCTGTTGAAGGGCTCTCTCAGATGATGAGATGACTTTTCCACCAACAC 763
Db 221 ProGlyGlyPheLeuPheGluGlyLeuSerAspGluAspPheHisProAsnThr 240

QY 764 AGGTCCACACCTCTAGCAGTACTCCAGCTCCCGCCAGCTCCCTCGGGTACAGTGA 823
Db 241 ArgSerThrProSerSerSerThrProSerSerArgProAlaSerLeuGlyTyrSerGly 260
QY 824 GCTGCTGGGCCCCGCCATCACCAGAGTAGCTGCCACCCCTTGGCCCTGGCCAGC 883
Db 261 AlaAlaGlyProArgProIleThrGlnSerGluLeuAlaThrAlaLeuAlaLeuAlaSer 280
QY 884 ACTCCGAGAGAGCTCTCACACGACTCTCTGGGACCCAGGCTTCTCTCAGGGACC 943
Db 281 ThrProGluSerSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
QY 944 TCACCAATGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
Db 301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAsnAspLeuPheSerGln 320
QY 1004 GCCTTACAGCATGCCCTTACAGGCTCTGGGACCCAGCTTCCAGAGCCAGTGGAGCC 1063
Db 321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnTyrGlnPro 340
QY 1064 CAGCTGCAGCAGCTGACATGGGATGGGATGGGATGGGATGGGATGGGATGGGATGGG 1122
Db 341 GlnLeuGlnGlnLeuArgAspMetGlyLeuGlnAspGluLeuSerLeuArgAlaLeu 360
QY 1123 CAGGCCACCCGTGGGACATCCCAAGCAGCCCTCGAGCTCATCTTTGCTGAGGAGCCCA 1182
Db 361 GlnAlaThrGlyGlyAspIleGlnAlaLeuGluLeuIlePheAlaGlyGlyAlaPro 380
RESULT 2
Q96S82 PRELIMINARY; PRT; 380 AA.
ID Q96S82;
AC Q96S82;
DT 01-DEC-2001 (TREXBurel. 19, Created);
DT 01-DEC-2001 (TREXBurel. 19, Last sequence update);
DT 01-MAR-2003 (TREXBurel. 23, Last annotation update);
DE Ubiquitin-like protein SBI32.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang W., Li N., Wan T., Cao X.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RS EMBL; AY037166; AK67643.1;
DR InterPro; IPR001901; SECE.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 1.
DR SMART; SM00213; UBQ; 1.
DR PROSITE; PS01367; SECE_SEC61G; 1.
DR PROSITE; PS00553; UBIQUITIN_2; 1.
SQ SEQUENCE 380 AA; 40581 MW; SABFA414F61F996C CRC64;
Alignment Scores:
Pred. No.: 1.67e-138 Length: 380
Score: 1904.00 Matches: 377
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 3
Query Match: 82.71% Indels: 1
DB: 4 Gaps: 0
US-09-735-251-3 (1-1273) x Q96S82 (1-380)
QY 44 ATGCTCTCTCAGACTGGCAGCTGGCGTGAATCGGTGACGACCCACTTACTCCAAAG 103
Db 1 MetSerLeuSerAspThrHisLeuAlaValLeuAlaAspGlnProLeuThrProlys 20
QY 104 TCTATCTCTCGGTGGCCAGACAGAACTGGGAGATATCTCGTACGGGCTATATATT 163

Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L., Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S., Hayashizaki Y.
"Functional annotation of a full-length mouse cDNA collection.";
Nature 409:685-690(2001).
EMBL: AK090025; RAH26033.1;
MCD: MGI:1916709; Z300004C1SRik.
InterPro: IPR023901; SecE.
InterPro: IPR030443; UBA domain.
InterPro: IPR033628; Ubiquitin.
Pfam: PF00627; UBA.1.
Pfam: PF00240; ubiquitin.1.
SMART: SM00165; UBA.1.
SMART: SM00213; UBG.1.
PROSITE: PS01067; SEC_ESEC6IG.1.
PROSITE: PS50053; UBIOUITIN.1.
SEQUENCE 380 AA; 40440 MW; OAF7DB0239BF7D11 CRC64;

Alignment Scores:
Pred. No.: 4,16e+135 Length: 380
Score: 1860.00 Matches: 367
Percent Similarity: 97.89% Conservative: 5
Best Local Similarity: 96.58% Mismatches: 8
Query Match: 80.80% Indels: 1
DB: 11 Gaps: 0

US-09-7355-251-3 (1-1273) x Q9D7P5 (1-380)
QY 44 ATGTCTTCTCCAGACTGGCACTGGCGGTGACACTGGCTGCAGACCCTTACTCCAAG 103
Db 1 MetSerLeuSerAsnTrpHisLeuAlaValLysLeuAlaAspGlnProLeuAlaProLys 20
QY 104 TCTATTCTTCGGTTGCCAGAGACAGAACTGGGAGAATACTCGCTAGGGGGCTATAGTATT 163
Db 21 SerIleuGlnLeuProGluThrGluLeuGlyGluTyrSerProGlyGlyTyrSerIle 40
QY 164 TCATTCTGAAGCAGCTTATTGTGTGGCAAACCTCCAGAGTGCTGTTCAGACCCTCAGCTG 223
Db 41 SerPheLeuLysGlnLeuIleAlaGlyLysLeuGlnGluSerValProAspProGluLeu 60
QY 224 ATTGATTTGATCTACTGTGGTCGGAGCTAAAGATGACAGACACTTCACCTTCTATGCG 283
Db 61 IleAspLeuIleTyrCysGlyArgLysLeuLysAspGlnThrLeuAspPheTyrGly 80
QY 284 ATTCAAATGGAGTCCACTGTCCATGTCTTCGAAAGCTCTCGCTCGAACCCTCATCAGAAA 343
Db 81 IleGlnProGlySerThrValHisValLeuArgLysSerTyrProGluProAspGlnLys 100
QY 344 CCGAAGCTGTGGNCAANGTGGCTGCATGAGAGAGTTCGGGTGTTCACACTGCCCTG 403
Db 101 ProGlnProValAspLysValAlaLeuArgGluPheArgValLeuHisThrAlaLeu 120
QY 404 CACAGCACGTCCTCTTACAGGGAGGCGGTCTTTTAAGATGCTCAGCAATAAGGAGTCTCTG 463
Db 121 HisSerSerSerTyrArgGluAlaValPheLysMetLeuSerAsnLysGluSerLeu 140
QY 464 GATCAGATTCATTGTGGCCACCCAGGCGCTCAGCAGTGACCTTATTCCTCTGGGGTTC 523
Db 141 AspGlnIleValAlaThrProGlyLeuSerSerAspProIleAlaLeuGlyValLeu 160
QY 524 CAGCAGAGGACCTCTTCTGTCTGCTGCTGATGCCCAATATGCTGATACGTGTGGCT 583
Db 161 GlnAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspThrLeuValPro 180
QY 584 GCTCACCAGCCCTCGTCAATGCGATTGCTCGTTCGCTCCTCCTAGCAGCAGTGCC 643
Db 181 AlaHisProAlaLeuValAsnAlaIleLeuValIleHisSerValAlaGlySerThr 200
QY 644 CCATGCTGGAGCTACTCTCTTCCGGAGGATCGCTCCAGCTCATACCGGATATG 703
Db 201 ProMetProGlyAlaAspSerSerArgSerMetProSerSerTyrArgAspMet 220
QY 704 CCAGTGGCTCTCTGTTTGAAGGGCTCTTCAGATGATGAGATGACTTTTCACCAACAC 763


```

Db      221 ProGlyGlyPheLeuPheAspAlaLeuSerAspGluAspPheHisProSerThr 240
QY      764 AGTCCACACCTCTAGAGTACTCCAGGCTCCGAGGCTCCCTGGGTACAGTGA 823
Db      241 ArgSerThrProSerSerThrProSerArgProAlaSerLeuGlyTyrSerGly 260
QY      824 GTGTGGGCGCGCCGATCAGAGTGAAGTGGGCAAGGCTTGGGCTGGGCGAG 883
Db      261 AlaAlaGlyProArgProIleThrGlnSerGluLeuAlaAlaLeuAlaSer 280
QY      884 ACTCCGAGAGCAGCTCTCAGACACCGACTCTGACACCCAGGTCATCTCTAGGAGC 943
Db      281 ThrProGluSerSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
QY      944 TCACCAATGCTCTGTGGTCCAGTCAGGAGAGCCCATCACAATGATCTCTTACGCAA 1003
Db      301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAsnAspLeuPheSerGln 320
QY      1004 GCCTACAGCATGCCCTTCAGGCTCTGGGCAAGCCAGCTTCAGAGCCAGTGGCAGCC 1063
Db      321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyHisProSerLeuGlnIleGlnTrpGlnPro 340
QY      1064 CAGCTCAGCAGCTAGTACATGGGCATCCAGAGCAGTACAGCTGAGCTCGGCC-CTG 1122
Db      341 GlnLeuGlnGlnLeuArgAspMetGlyTleGlnAspAspGlnLeuSerLeuArgAlaLeu 360
QY      1123 CAGGCCACCGGTGGGACATCCAAAGCAGCCCGGAGCTCATCTTGTGGAGGAGCCCA 1182
Db      361 GlnAlaThrGlyGlyAspIleGlnAlaAlaLeuGluLeuIlePheAlaGlyAlaPro 380

RESULT 5
ID Q8K141 PRELIMINARY; PRO: 554 AA.
AC Q8K141
DI 01-OCT-2002 (TrEMBLrel. 22, Created:
DI 01-OCT-2002 (TrEMBLrel. 22, Last sequence update:
DI 01-MAR-2003 (TrEMBLrel. 23, Last annotation update:
DE Hypothetical protein.
CS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10390;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Strausberg R.;
RL EMBL: BC028857; AAH28857.1;
DR InterPro; IPR000449; UBA_domain.
DR Pfam; PF00627; UBA; 1.
DR SMART; SMC0165; UBA; 1.
DR SMART; SMC0213; UBO; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.
SQ SEQUENCE 554 AA; 58677 MW; F9255C9985A4C029 CR054;

Alignment Scores:
Pred. No.: 4,57e-06 Length: 554
Score: 188.50 Matches: 104
Percent Similarity: 39.74% Conservative: 75
Best Local Similarity: 22.96% Mismatches: 174
Query Match: 8.19% Indels: 99
DB: 11 Gaps: 21

US-09-735-251-3 (1-1273) x Q8K141 (1-554)
QY 86 CAGCAGCTTACTCCAGAGTCTATCTCTGCTTCCAGAGACAGACTGGGAGATATCG 145
Db 24 GluProIleMetLysValThrValLysThrProLys---GluLysGluGluPheAla 42
QY 146 CTA---GGGGGCTATGATTTCATTTC---GAACAGCTATTCTGCTGGCAATCCAGGAG 202

```

```

Db      43 ValProGluAsnSerSerValGlnGlnPheLysGluGluIleSerLysArgPheLysSer 62
QY      203 TCTGTTCAGACCCCTGAGCTGATCTGATCTACTGTCTGCGAAGCTTAAAGATGAC 262
Db      63 HisIle-----AspGlnLeuValLeuIlePheAlaGlyLysIleLeuLysAspGln 79
QY      263 CAGACACTTGACTCTTATGGCATTCACCTGGTCCACTGCTCCATGTTCTCGGAAG--- 319
Db      80 AspThrLeuSerGlnHisGlyIleHisaspGlyLeuThrValHisLeuValIleLysThr 99
QY      319 ----- 319
Db      100 GlnAsnAspProGlnAspAsnSerAlaGlnGlnThrAsnAlaProGlySerThrValThr 119
QY      320 TCTGTGGCTGACCTGATCAGAAACCGGAACTGTGGACAAGTGGCTGCCATCAGAGAG 379
Db      120 SerSerProAlaProAspSer-----AsnProThrSerGlySerAlaAlaAsnSerSer 137
QY      380 TTCCGGGTG-----TTGCACACTGCCCCTGCACAGCAGCTCC 415
Db      138 PheGlyValGlyGlyLeuGlyLeuAlaGlyLeuSerSerLeuGlyLeuAsnThrThr 157
QY      416 TCTTACAGGAGCGGTCTTTAAGATG-----CTCAGCAATAAGAGTCTCTG 463
Db      158 AsnPheSerGluLeuGlnSerGlnMetGlnArgGlnLeuLeuSerAsnProGluMetMet 177
QY      464 GATCAGTCATTTGTGGCCACCCAGCCCTCAGCAGTGCACCTATTGCTTCTTGGGTTCTC 523
Db      178 ValGlnIleMet-----GluAsnProPheValGlnSerMetLeu 190
QY      524 CAGCAAGAGCTCTTCTCT-----GTCTTCGCTGATCCCAATATGCTTGATAGTTG 577
Db      191 SerAspProAspLeuMetArgGlnLeuIleMetAlaAsnProGlnMet---GlnGlnLeu 209
QY      578 GTGCTGCTCACCAGCCCTCTCAATGCC-----ATTGCTGCTGTTCTG 622
Db      210 IleGlnArgAsnProIleSerHisMetLeuAsnAsnProAspIleMetArgGlnThr 229
QY      623 CACTCTGTAGCAGGAGCCCAATGCTGGAGCTGACTCTCTTCCCGGAGCATGCC 682
Db      230 LeuGluLeuAlaArgAsnProAlaMetMetGlnMetMetArgGlnAspArgAla 249
QY      683 TCCAGCTCATACCGGGATATGCCAGTGGCTTC-----CTGTTTGAAGGG 727
Db      250 LeuSerAsnLeuGluSerIleProGlyGlyTyrAsnAlaLeuArgArgMetFyrThrAsp 269
QY      728 CTCTCAGATGAT-----GAGGATGACTTTTACCCCAACACACAGGTCACCA 772
Db      270 IleGlnGluProMetLeuAsnAlaAlaGlnGlnPheGlyGlyAsn----- 285
QY      773 CCTCTAGCAGTACTCCAGCTCCCGCCAGCTCTCCCTGGGGTACAGTGGAGCTCTCTGG 832
Db      286 ProPheAlaSerLeuValSerSerSer-----SerSerAlaGluGly 299
QY      833 CCGCGGCCCATCACCAGAGTGCAGTGGCCCGCTTGGCCCTGGCCAGCTCCGAG 892
Db      300 ThrGlnPro---SerArgThrGluAsnArgAspProLeuProAsnProThrAlaProGln 318
QY      893 AGCAGCTCTCACACACCGAGCTCTCTGACCCAGGCTCATCTCTCAGGAGCTCCACCAATG 952
Db      319 ThrSerGlnSerSerProAlaSerGlyThrThrGlySerThrThrAsnThrMetSerThr 338
QY      953 TCTCTGTGTCAGTCCAGGAGGCGCCATCACCATGATCTCTTCCAGCAAGCCCTACAG 1012
Db      339 SerGlyGlyThrAlaThrSerThrProAlaGlyGlnSerThrSerGlyProSerLeuVal 358
QY      1013 CATGCCCTTCAGGCCCTCT-----GGGACGCCAGCTTCAGAGCAGTGGGAG----- 1060
Db      359 ProGlyValAlaSerMetPheAsnThrProGlyMetGlnSerLeuLeuGlnIle 378
QY      1061 -----CCCAGCTGCAGCAGCTACTGTGACATGGGCATCCAGAGCATGAGCTGAGC 1111

```


QY 1214 AGTTCAGAGC 1225

Db 454 SerGlyAlaGly 457

RESULT 7

Q8C835 PRELIMINARY; FRT: 554 AA.
AC Q8C835;
D7 01-MAR-2003 (TEMBUREL 23, Created;
D7 01-MAR-2003 (TEMBUREL 23, Last sequence update);
D7 01-MAR-2003 (TEMBUREL 23, Last annotation update);
DE DA4: homolog;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466831;
RA The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase 1 & 2 Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
60,770 full-length cDNAs";
RL Nature 420:563-573 (2002).
DR EMBL: AK048534; BAC33365.1;
SQ SEQUENCE 554 AA; 58605 MW; 2FC46079D945A29 CRC64;

Alignment Scores:

Pred. No.: 1,45e-05 Length: 554
Score: 182.00 Matches: 125
Percent Similarity: 34.96% Conservative: 68
Best Local Similarity: 22.64% Mismatches: 151
Query Match: 7.91% Indels: 209
Gaps: 26

US-09-735-251-3 (1-1273) x Q8C835 (1-554):

QY 86 CAGCCACTACTCCAAAGCTATTCTGGTGGCCAGACAGAACTGGAGAACTACTG 145
Db 454 SerGlyAlaGly 457
24 GluProLysIleMetLysValThrValLysThrProLys---GluLysGluLysPheAla 42
146 CTA---GGGGGCTATAGTATTCTGCTGCAACGACTATTGCTGGCAAACTCCAGAG 202
Db 43 ValProGluAsnSerSerValGlnHisMetLysHisGluLysSerLysPheLysSer 62
203 TCTGTCACAGACCTGAGCTGATATATGATATGCTGCTGCTGCTGCTGCTGCTGCTG 262
Db 63 HisLe---AspGlnLeuValLysLysPheAaGlyLysIleLeuLysAspGln 79
263 CAGACACTGCTTCTATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
Db 80 AspThrLeuSerGlnHisGlyLysHisAspGlyLeuThrValHisLeuValLysLys 99
319
Db 100 GluAsnArgProGlnAspAsnSerAlaGlnGlnThrAsnAlaProGlySerThrValThr 119
320 TCTTGGCTGAACTGATCAGAAACCGGAACCTGTGGACAAAGTGGCTGCTGCTGCTG 379
Db 120 SerSerProAlaProAspSer---AsnProThrSerGlySerAlaAlaAsnSerSer 137
380 TTCGGGGT---ACCTCTCTCTACAGGGAGGCG---TTCGACACTGCC 400
Db 138 PheGlyValGlyLeuGlyGlyLeuAaGlyLeuSerSerLeuGlyLeuAsnThrThr 157
401 CTGCACAGC---ACCTCTCTCTACAGGGAGGCG---TTCGACACTGCC 430
Db 158 AsnPheSerGluLeuGlnSerGlnMetGlnArgGlnLeuLeuSerAsnProGlyMetMet 177
431
Db 178 ValGlnIleMetGluAsnProPheValGlnSerMetLeuSerAsnProAspLeuMetArg 197

QY 467 CAGATCATTTGGSCACCCAGGCTCAGCACT-----GACCCTATTCTCTT 514
Db 198 GlnLeuIleMetAlaAsnProGlnMetGlnGlnLeuLeuGlnArgAsnProGluLeuSer 217
QY 515 GGGGTTCTCCAGCAAGGACCTCTTCTGCTTGGCTGATCCCAATATGCTGATACG 574
Db 218 HisMetLeuAsnAsnProAspIle-----MetArgGlnThr 229
QY 575 TTG---GTSCCTGCTCACCAGCCCTGCTCAATGCCATTGCTCTGTTCTG---CACTCC 628
Db 230 LeuGluLeuAlaArgAsnProAlaMetMetGlnGluMetMetArgAsnGlnAspArgAla 249
QY 629 GTAGCAGGAGTGGCCCAATGGCTGACTACTCTTCTCCGAGAGTACCTCCCTCCAGC 688
Db 250 LeuSerAsnLeuGluSerIleProGlyGlyTyrAsnAlaLeuArgMet----- 266
QY 689 TCATACGGGATATGCCA-----GGTGGCTTC 715
Db 267 ---TyrThrAspIleGlnGluProMetLeuAsnAlaGlnGlnPheGlyGlyAsn 285
QY 716 CTGTTGAAGGGCTC-----TCA 733
Db 286 PropheAlaSerLeuValSerSerSerSerSerAlaGluGlyThrGlnProSerArgThr 305
QY 734 GATGATGAGGATGACTTTCCACCAAC-----ACCAGGTCCACACCC 775
Db 306 GluAsnArgAspProLeu---ProAsnProTAlaProGlnThrSerGlnSerSerPro 324
QY 776 TTAGCAGTACTCCAGCTCCGCCAGCTCCCTGGGGTACAGTGA----- 823
Db 325 AlaSerGlyThrThrGlySerThrThrAsnThrMetSerThrSerGlyThrAlaThr 344
QY 824 -----GCTGCTGGGCC----- 835
Db 345 SerThrProAlaGlyGlnSerThrSerGlyProSerLeuValProGlyAlaGlyAlaSer 364
QY 836 -----CGGCCCATCACC 847
Db 365 MetPheAsnThrProGlyMetGlnSerLeuLeuGlnGlnIleThrGluAsnProGlnLeu 384
QY 848 CAGAGTGAGTGGCCACCCCTTGGCCCTGGCCAGCACTCCGAGAGGAGCTCTCACACA 907
Db 385 MetGlnAsnMetLeuSerAlaProTyrMetArgSerMetLeuGlnSerLeuSerGlnAsn 404
QY 908 CCG----- 910
Db 405 ProAspLeuAlaAlaGlnMetGlnAsnProAspThrLeuSerAlaMetSerAsnProArg 424
QY 910 ----- 910
Db 425 AlaMetGlnAlaLeuLeuGlnIleGlnGlnGlyLeuGlnThrLeuAlaThrGluAlaPro 444
QY 911 -----ACTCCT-----GGCACCAGCGGTCAATTCCTCAGCG 940
Db 445 GlyLeuIleProGlyPheThrProGlyLeuAlaAlaGlyAsnSerGlyGlySerSerGly 464
QY 941 ACTCCTCAATATGCC-----TCTGGTGTCCAGTCAAGGAGCCCATC 982
Db 465 ThrAsnAlaProSerThrAlaProSerGluAspThrAsnProGlnGlyThr----- 482
QY 983 ACCAATCATCTTTCAGCCAGCCCTACAGCATGCCCTTCAGGCTCTGGG-----CAG 1036
Db 483 AlaGluProGlyHisGlnGlnPheIleGlnGlnMetLeuGlnAlaLeuAlaGlyValAsn 502
QY 1037 CCCAGCCTTCAGAGC-----CAGTGGCAGCCCGCAGCTGACAGCTGCTACATG 1087
Db 503 ProGlnLeuGlnSerProGluValArgPheGlnGlnGlnLeuGlnLeuSerAlaMet 522
QY 1088 GGATCCAGGACCATGAGCTGAGCTGCG---GGCTGCGAGCCCGCTGGGAGCATCCAA 1146
Db 523 GlyPheLeuAsnArgGluAlaAsnLeuGlnAlaLeuAlaThrGlyGlyAspIleAsn 542

QY 1147 GCAGCCCTGGAGCTCATCTTCTGGAGGAGCCCA 1162
 DB 1147 GCAGCCCTGGAGCTCATCTTCTGGAGGAGCCCA 1162
 543 AialaLeuGluArgLeu---GlySerGlnPro 553

RESULT 8

Q8R317 PRELIMINARY; PRT: 582 AA.
 AC Q8R317;
 DT 01-JUN-2002 (TREMBlrel. 21, Created)
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Similar to ubiquitin 1.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC026847; AAH26847.1;
 DR InterPro; IPR006636; STILL;
 DR InterPro; IPR000449; UBA domain;
 DR InterPro; IPR00626; Ubiquitin;
 DR Pfam; PF00627; UBA; 1;
 DR Pfam; PF0245; ubiquitin; 1;
 DR SMART; SM00727; STILL; 4;
 DR SMART; SM00165; UBA; 1;
 DR SMART; SM00213; UBO; 1;
 DR PROSITE; PS53053; UBQUITIN 2; 1;
 SC SEQUENCE 582 AA; 61976 MW; E0CF987947C3A4 CRC64;

Alignment Scores:

Pred. No.: 2,74e-05 Length: 582
 Score: 178.50 Matches: 91
 Percent Similarity: 40.21% Conservative: 65
 Best Local Similarity: 23.45% Mismatches: 149
 Query Match: 7.75% Indels: 93
 Gaps: 11

US-09-735-251-3 (1-1273) x Q8R317 (1-582)

QY 86 CAGTCCTACTCCAAAGTCTATTCTTGGTGTCCAGACAGACAACTGGGAGAACTACG 145
 DB 86 CAGTCCTACTCCAAAGTCTATTCTTGGTGTCCAGACAGACAACTGGGAGAACTACG 145
 24 GluProLysLeuMetLysValThrValLysThrProLysLeuLysGluGluProLysVal 42
 QY 146 CTA GGGGCTATATATATTTTAAATATATATATATATATATATATATATATATAT 202
 DB 43 ValProGluAsnSerSerValGlnSerProLysLeuLysLeuLysLeuLysLeuLysLeu 62
 QY 203 TCTTTCAGACCTGACGCTGATGATCTGATCTGATCTGATCTGATCTGATCTGATCTGAT 262
 DB 63 Histile-----AspGlnLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 79
 QY 263 CAGACCTTGACTTCTATGGGATTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
 DB 80 AspThrLeuSerGlnHisGlyLeuHisAspGlyLeuThrValHisLeuValLeuLysThr 99
 QY 319 ----- 319
 DB 100 GlnAsnArgProGlnAspAsnSerAlaGlnGlnThrAsnAlaProGlySerThrValThr 119
 QY 320 TCTGGCTGACCTGATCAGAACCGGAGACCTGGACAAATGGCTGCCATGAGAGAG 379
 DB 120 SerSerProAlaProAspSer-----AsnProThrSerGlySerAlaAlaAsnSerSer 137
 QY 380 TTCGGGTG-----TTGCACACTGCCCTGCACAGCAGCTCC 415
 DB 138 PheGlyValGlyGlyLeuGlyGlyLeuAlaGlyLeuSerSerLeuGlyLeuAsnThrThr 157
 QY 416 TCTTACAGGAGGGCGGTCTTAAAGATG-----CTCAGCAATAAGGAGCTCTCG 463
 DB 158 AsnPheSerGluLeuGlnSerGlnMetGlnArgGlnLeuLeuSerAsnProGluMetMet 177

QY 464 GATCAGATCATTTGGCCACCCAGCCCTCAGAGTGAACCTATTGCTCTTGGGCTCTC 523
 DB 178 ValGlnLeuMet-----GluAsnProPheValGlnSerMetLeu 190
 QY 524 CAGGACAAAGACCTCTTCTCT-----CTCTTCGCTGATCCCAATATGCTGTATACGTTG 577
 DB 191 SerAsnProAspLeuMetArgGlnLeuLeuMetAlaAsnProGlnMet---GlnGlnLeu 209
 QY 578 GTGCTGTCTCACCAGCCCTCGTCAATGCC-----ATTGCTCTGTTCTG 622
 DB 210 IleGlnArgAsnProGluLeuSerHisMetLeuAsnAsnProAspAlleMetArgGlnThr 229
 QY 623 CACTCCGAGGAGCAGTGCCTCCCAATGCTGGACTGACTCTCTTCCGAGCATGCC 682
 DB 230 LeuGluLeuAlaAaGAsnProAlaMetMetGlnGluMetMetArgGlnAsnArgAla 249
 QY 683 TCAGCTCATACCGGGATATCCAGGTGGCTTC-----CTGTTTGAAGGG 727
 DB 250 LeuSerAsnLeuGluSerIleProGlyGlyTyrAsnAlaLeuArgArgMetTyrThrAsp 269
 QY 728 CTCTCAGATGAT-----CAGGATGACTTTCACCCAAACACACAGGTCACCA 772
 DB 270 IleGlnGlnProMetLeuAsnAlaAlaGlnGlnPheGlyGlyAsn----- 285
 QY 773 CCTCTAGCAGTACTCCAGCTCCCGCCAGCCTCCCTGGGGTACAGTGGAGCTGCTGGG 832
 DB 286 ProPheAlaSerLeuValSerSer-----SerSerAlaGluGly 299
 QY 833 CCCCAGGCTCATCCAGCAGTGAAGTGCACCGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 892
 DB 300 ThrGlnPro---SerArgThrGluAsnArgAspProLeuProAsnProTrpAlaProGln 318
 QY 893 AGCAGCTCTCACACACCCAGCTCCCTGGCAGCCAGGCTCATCTCAGGAGCCTCACCAGT 952
 DB 319 ThrSerGlnSerSerProAlaSerGlyThrThrGlySerThrThrAsnThrMetSerThr 338
 QY 953 TCCTCTGTGTCTCAGTCCAGGAGCCCATCAATGATCTCTTCCAGCAAGCCCTCAG 1012
 DB 339 SerGlyGlyThrAlaThrSerThrProAlaGlyGlnSerThrSerGlyProSerLeuVal 358
 QY 1013 CATGCGCTCTCAGGCTCT-----GGGAGCCAGCCTTACAGCAGCAGTGGCAG----- 1060
 DB 359 ProGlyValAlaSerMetPheAsnThrProGlyMetGlnSerLeuLeuGlnGlnIle 378
 QY 1061 -----CCCAGCTCCAGCAG 1075
 DB 379 ThrGlnAsnProGlnLeuMetGln 385

RESULT 9

Q8C7T4 PRELIMINARY; PRT: 555 AA.
 ID Q8C7T4;
 AC Q8C7T4;
 DT 01-MAR-2003 (TREMBlrel. 23, Created)
 DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE DA41 homolog.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 DR EMBL; AK049298; BAC33666.1; --
 SC SEQUENCE 555 AA; 59497 MW; 7275E3735A039375 CRC64;

UY 323 GAA

QY	716	-----CGTTTGAAGGGCTCTCAGATGAT-----	-----GAGGATGACCTTT	751
Db	279	LeuArgArgMetTyrThrAspIleGlnGluProMetLeuAsnAlaAlaGlnGluInPhe		298
QY	752	CACCCAAACACCGGTCCACACCCCTCTAGCAGACTCCCGAGCTCCCGCCGAGCTCCCTG		811
Db	299	GlyGlyAsnProPheAlaThrValGlySerSerThrSerGly-----		313
QY	812	GGGTACAGTGAAGCTGTGGCGCCCGCCCATCACCCAGAGTGAAGCTGGCCACCGCTTG		871

[illegible]

872	GCCTGGCCAGCACTCGGAGAGAGAGCTCTCACACACCGACTCTCTGGCACCACCGAGGTCAAT	931
QY		
332	AlaProL-oProThrThrGlnThrAlaAlaThrThrThrThrThrThrThrThrThrThrSerSer	351
Db		
932	TCCTCAGGAGACTCACCAATGTCTCTGGTGTCAGTCAGGG	973
QY		

[illegible]

	974	----	ACCCCATCACCAGTATCTCTCAGC-----	101
Dc	372	AsnTyValAlaSerIlePheSerThrProGlyMetGlnSerLeuGlnGlnIleThr	391	
Qy	1001	-----CAAGGCCCTACAGCATGCCTCTCAGCGCTCTGGCGAGCCAGCCTTCAGAGC	105	
Dc	392	GluAsnProGlnLeuIleGlnAsnMetLeuSerAla-----ProTyrMetArgSer	408	
Qy	1052	CAGTGGCAGCCCCCAGCTCGCAGCAGCTACGTGACATGGCGCATCCAGGACCATGATGCTGAGC	111	
Dc	409	MetMetG:n---SerLeuSerGlnAsnProAspMetAlaAlaGlnMetMetLeuSerSer	427	
Qy	1112	----CTGCGGCCCTGCAGGCCACCGTGCGGACATCCAAAGCAGCGCTGGAGCTCATCTTTG	116	
Dc	428	ProLeuThetThrSerAsnProGlnLeuGlnGlnGlnMetArgProGlnLeuProAsnPhe	445	
Qy	1169	CTGAGGAGGAG-----CCCATGAATCCCTGCTCCCTCAACCCCGCAGCAAGTTGC	121	
Dc	448	LeuGlnG:nMetG:nAsnProGluThrIleAlaAlaMetSerAsnProArgAla-----	465	
Qy	1220	AGAGCGCTACTCCCTTGGGAGGCACTCATGAAGTGCTGCCATCTCTC	1267	
Dc	466	-----MetGlnAlaLeuMetGlnIleGlnGlnGlyLeu	476	
 RESULT 13 Q3SN46				
ID	Q3SN46	PREDICINARY;	PRT; 839 AA.	
DT	C1-MAY-2003	(TREMBLrel. 13, Created)		
DT	C1-MAY-2000	(TREMBLrel. 13, last sequence update)		
DT	C1-MAR-2003	(TREMBLrel. 23, last annotation update)		
DE	Extensin-like protein.			
GN	F28A21.80 OR A14G18670.			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OX	eurosid II; Brassicales; Brassicaceae; Arabidopsi.			
NCBI_TaxID=3702;				
RP	SEQUENCE FROM N.A.			
RA	Bevan M., Mueller M.W., Muendlein A., Felber R., Bancroft I.,			
RA	Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.			
[2]				
RP	SEQUENCE FROM N.A.			
RA	EU Arabidopsis sequencing project;			
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.			
[3]				
RP	SEQUENCE FROM N.A.			
RA	Mueller M.W., Muendlein A., Felber R., Mewes H.W., Lemcke K.,			
RA	Mayer K.F.X.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			

Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]

RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

CR EMBL; AL035526; CAB37452.1; ...
 DR EMBL; AL161549; CAB78869.1; ...
 DR InterPro; IPR002951; Atrophidin.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007290; LRR_P1ant.
 DR InterPro; IPR002965; P-rich_extens.
 CR Pfam; PF00560; LRR; 4.
 DR PRINTS; PR01222; ATROPHIN.
 DR PRINTS; PR01217; PRICHEXTENSK.
 DR PROSITE; PS50502; LRR_PS; 3.
 SQ SEQUENCE 839 AA: 90243 MW: 38404 kDa; 29755 CRC64;

Alignment Scores:
 Pred. No.: 0.00306 Length: 839
 Score: 165.50 Matches: 97
 Percent Similarity: 34.96% Conservative: 25
 Best Local Similarity: 27.71% Mismatches: 121
 Query Match: 7.13% Indels: 107
 DB: 10 Gaps: 14

US-09-735-251-3 (1-1273) x Q9JUP9 (1-839)

QY 312 TCGAAGCTCTGGCCCTG...AACCTGATCAGAAACCGGAACCTGTGGACAAAGTGGCTG 368
 DB 337 CysGlnLeuProArgLeuGluAsnProThrPheSerTyrAsnProPheThrGly----- 354
 QY 369 CCAATAGAGATTTCGGGTGTGGACATGGCT-----TUCACA 407
 DB 355 -----GluProProValCysLeuGlyLeuProGlyPheAspArgAtGAsnCysLeu 372
 QY 458 GCAGCTCTCTTACAGGAGGAGTGTCTTAAGATGCTCAGCAATAAGAGTCTCTGGATC 467
 DB 373 ProAaArgProAlaGlnArgSerProGlyCysAlaAlaPheSerSerLeuProPr 392
 QY 468 AGATCATTTGGCCAC-----GCCAG 488
 DB 392 eValAspCysGlySerPheGlyCysGlyArgSerThrArgProProValValProSe 412
 QY 489 GCTCAGAGTACGCTTATCTCTGGGTCTCCAGGACAGGACCTCTCTCTCTCT 548
 DB 412 rProProThrPro-----SerProGlyGlySerPro----- 423
 QY 549 TCGCTATCCCAATATGCTTATATAGTATGATGATGATGATGATGATGATGATGATGAT 609
 DB 424 -----SerProSerSerSerSerSerSerSerSerSerSerSerSerSerSer 430
 QY 609 TTGCTCTGGTCTGCACTCCCTAGAGAGTATGATGATGATGATGATGATGATGATGATGAT 658
 DB 430 oSerPro-Proile-----ThrValProSerProProProProProProProProPro 444
 QY 669 CCGGAGCATGCCCTCCAGCTCAGACAGAGATATGCGGAGGTGTTCTCTTTGAGGGG 728
 DB 445 ProGlyGlySerProProSerProSerSerSerSerSerSerSerSerSerSerSer 455
 QY 729 TCTCAGATGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788
 DB 456 -----ProSerProProSerProProProProProProProProProProProPro 462
 QY 789 CCAGCTCCGCCACCTCCCTGGGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 848
 DB 463 ProSerProGlySerProProProProProProProProProProProProProProPro 481
 QY 849 AGAGTACGATGCGCCACCGCTTGGCCCTGGCCAGCACTCCGAGAGAGTCTCTCACAC 908
 DB 482 ProSerSerProThrThrProThrPro--GlyGlySerProProSerSerProThrThr 501
 QY 909 CCACTCTCGGACCCAGGCTCATCTCTCAGGAGTCTACCAATGCTCTCTGCTGCTCCAGT 969
 DB 501 roThrProGlyGlySerProProSerSerProThrThrProSerProGlyGly-SerPro 520

QY 969 CAGGACGCCCATCAATGATCTCTTACAGCAAGCCCTACAGCATGCCCTTCAAGCCT 1028
 DB 521 -----ProSerProSerProSerProSerProSerProSerProSerProSerPro 536
 QY 1029 CTGGGACGCCCATGCTTACAGCAGTGGCAGCCCGAGCTGCGAGCAGCTAGCTGACATGG 1088
 DB 537 ProSerThrProThr-----SerProGlySerProProSerProSerProSerPro 554
 QY 1089 GCATCCAGGACGATGAGCTGAGCTGCGGCCCTGCGAGCCCGGTCGGGACATCCAGC 1148
 DB 555 SerSerPro:leProSer-ProProThr-ProSerThrProProThrProleSerPro 574
 QY 1149 AGCCCTGAGTCATCTTTGCTGGAGGAGCC-----CATGAACCTCCCTGCTT 1196
 DB 574 LYG:nAsnSerProProle:leProSerProProPheThrGlyProSerProProSer 594
 QY 1197 CCCTGAGCCCGCAGCAAGTGGAGAGCTACTGCCCTTGGGAGGCACTCATGAGGTGC 1256
 DB 594 erProSerProPro-----Leup 600
 QY 1257 CTCCATCTCTCCCT 1270
 DB 600 roProVallePro 604

RESULT 14
 Q9JUP9 PRELIMINARY; PRT: 582 AA.
 AC Q9JUP9;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE DA41.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA Ozaki T., Hishiki T., Nakagawa A., Sakiyama S.;
 RT "Identification of a new protein that can associate with DAN.";
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL; S87950; BAA92267.1; ...
 DR HSSP; P02448; ITBE.
 DR InterPro; IPR006636; STIL.
 DR InterPro; IPR00449; UBA domain.
 DR InterPro; IPR006626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; ubiquitin; 1.
 DR SMART; SM00727; STIL; 4.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00213; UBQ; 1.
 DR PROSITE; PS50053; UBIQUITIN 2; 1.
 SQ SEQUENCE 582 AA: 62071 MW: 66515 kDa; 90971ESB CRC64;

Alignment Scores:
 Pred. No.: 0.00301 Length: 582
 Score: 165.00 Matches: 117
 Percent Similarity: 32.70% Conservative: 73
 Best Local Similarity: 20.14% Mismatches: 153
 Query Match: 7.17% Indels: 239
 DB: 11 Gaps: 24

US-09-735-251-3 (1-1273) x Q9JUP9 (1-582)

QY 86 CAGCCATCTTACTCCAACTCTTCTTGGTTCGACAGACAGCACTGGAGATCTCG 145
 DB 24 GluProLysMetLeuValThrValLysThrProLys--GlyLysGluGluPheAla 42
 QY 146 CTA---GGGGCTATAGTATTTCCTTCTGAGCAGCTTATCTGCGCAAACTCCAGGAG 202
 DB 43 ValProGluAsnSerSerValGlnGlnPheLysGluLeuLeuSerLysArgPheLys 62

RP SEQUENCE FROM N.A.
 RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Bach C., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kariya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
 RA Ecker J., Theologis A., Davis R.W.
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
 RA Palm C.J., Bowser L., Jones T., Bach C., Carninci P., Chen H.,
 RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kariya A., Kawai J.,
 RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
 RA Ecker J., Theologis A., Davis R.W.
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kariya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Pham P.K.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A., Jones T.
 RC "Arabidopsis Open Reading Frame (ORF) Clones".
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC007127; AAC25137.2;
 DR EMBL: AF360159; AAK25869.1;
 DR EMBL: AY042828; AAK68768.1;
 DR EMBL: AY081450; AAM10012.1;
 DR EMBL: AY113886; AAV44914.1;
 DR InterPro: IPR003449; UBA domain.
 DR InterPro: IPR00626; UBAquitin.
 DR Pfam: PF00627; UBA; 1.
 DR PROSITE: PS00553; UBIQUITIN 2;
 SQ SEQUENCE 538 AA; 57285 MW; AF2BFHF6413DE0 CRC64;

Alignment Scores:
 Pred. No.: 0.000322 Length: 538
 Score: 164.50 Matches: 99
 Percent Similarity: 35.44% Conservative: 40
 Best Local Similarity: 23.94% Identical: 14
 Query Match: 7.15% Gaps: 14
 DB: 10

US-09-735-251-3 (1-1273) x 09S113 (1-534)

QY 158 AGTATTCATTCTGAAGCAGCCTATTGTTGGAAAGTTCAGAGCTCTTTCCAGACCT 217
 DB :|||||
 DB 39 ThrValGluSerPheLysGluLeuAla...GlnAsnSerValProAla 55
 QY 218 GAGCTGATTGATCTGCTACTCTGGTCAAGATTAAGAGATGACAGCACTTACTCT 277
 DB :|||||
 DB 56 AsnGlnGlnArgLeuLeuTyrLysGlyArgIleLeuLysAspAspGlnThrLeuLeuSer 75
 QY 278 TATGGCTTCAACCTGGCTGCATCTGCATCTTTTGGGAAGCTCCGGCTGAACCTGAT 337
 DB :|||||
 DB 76 TyrGlyLeuGlnAaAspHisThrValHisMetValArgGlyPheValProSerSerPro 95
 QY 338 CAGAAACCG----- 346
 DB :|||
 DB 96 SerAlaProAlaAlaAsnAaGlyAsnGlnThrThrAlaProGlnAlaValGlySerAsn 115
 QY 347 -----GACCTGTG 355
 DB :|||
 DB 116 AspSerSerAsnLeuGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 135
 QY 156 CACAAAGTGGCTGCATGAGAGAGTTC-----CGGTG 388
 DB :|||

Search completed: November 12, 2003, 06:49:21
 Job time : 75.5 secs

DB 136 GlyGlyGlyAsnAlaMetAlaGlyLeuPheGlySerGlyLeuProAspLeuGluGlnAla 155
 QY 389 TTGCACACTGCTCCCTGCACAGCAGCTCTCTTACAGGGAG-----CGGTG 433
 DB :|||
 DB 156 GlnGlnGlnLeuAlaGlnAsnProAsnMetIleArgGluMetMetAsnThrProAlaIle 175
 QY 434 TTAAAGATGTCAGCAATAAGGAGTCTCTGGATCAGATCATTTGTGCCACCCACGCTC 493
 DB :|||||
 DB 176 GlnAsnLeuMetAsnAsnProGluPheMetArgSerMetIleMetAsnAsnProGlnMet 195
 QY 494 AGCAGTACCCCTATTCTCTTGGGGTTCTCCAGGACAAG-----GACTCTTCTCTGTC 547
 DB :|||||
 DB 196 ArgGlu-----LeuValAspArgAsnProGluLeuGlyHisVal 208
 QY 548 TTGCTGATCCCAATATGCTT---GATAGTGTGGTGCCTGCT---CACCCAGCCCTGCTC 601
 DB :|||||
 DB 209 LeuAsnAspProSerIleLeuArgGlnThrLeuGluAlaAlaArgAsnProGluLeuMet 228
 QY 602 AATGCCATTGCTCTGCTGCTTCTGCACCTCCGTAGCAGCAGTGCCTCCCAATGCTGGACTGAC 661
 DB :|||
 DB 229 ArgGluMet-----MetArgAsnThrAsp 236
 QY 662 TCCTCTTCCCGAGAGTGCCTCCAGCTCATCCGGGATATCCAGGTGGCTTC----- 715
 DB :|||||
 DB 237 -----ArgAlaMet-----SerAsnIleGluSerMetProGluGlyPheAsnMet 251
 QY 716 -----CTGTTGAAGGGCTCTCAGATGATGAGTGTTCACCCCAACACCAGG 766
 DB :|||
 DB 252 LeuArgArgMetTyrGluAsnVal-----GlnGluProLeuMetAsnAlaThrThrMet 269
 QY 767 TCCACACCTCTTAGCAGTACTCCAGCTCCCGCCCA-----GCCTCCCTGGGTACAGT 820
 DB :|||
 DB 270 SerGluAsnAlaGlyAsnAsnThrSerSerAsnProPheAlaLeuLeuGlyAsnGln 289
 QY 821 GGAGCTGCTGGCGCCCGCCCATCACAGAGTGTGCTGCGCACCCAGGCTTCTCTCCAGG 880
 DB :|||||
 DB 290 GlyValThrThrGlnGlySerAspThrSerAsnAsnIleSerAlaProAsnAlaGluThr 309
 QY 881 AGCACTCCGAGGAGCAGCTCTCACACACCGAGCTCTGCGCACCCAGGCTTCTCTCCAGG 940
 DB :|||||
 DB 310 GlyThrProAsnAlaAsn-----ProLeuProAsnProThrGlyAlaThrAlaGly 326
 QY 941 -----ACCTCACCAATGTCTCTGCTGCTCCAGTACAGG----- 973
 DB :|||||
 DB 327 GlnThrThrAlaProGlyArgThrAsnAaGlyLeuGlyGlyLeuGlyGlyGlyGly 346
 QY 974 -----ACGCCATCACAAT 988
 DB :|||
 DB 347 LeuGlyGlyLeuGlyMetLeuGlyAlaAspSerProLeuGlyAlaThrProAspAlaSer 366
 QY 989 GATCTCTTCAGCAAGCCCTACAGCATGCTTTCAGGCTCTGGGAGCCGCGCTTCAG 1048
 DB :|||
 DB 367 GlnLeu---SerGlnIleLeuGlnAsnProAlaMetSerGlnMetMetGlnSerValLeu 385
 QY 1049 AGCCAGTGGCAG-----CCCCAGCTGCAGCAGCTACCT 1081
 DB :|||||
 DB 386 SerAsnProGlnTyrMetAsnGlnLeuMetSerLeuAsnProGlnLeuArgSerMetLeu 405
 QY 1082 GACATGGCATCCAGGACGATGAGCTGAGCTGCGGCCCC 1120
 DB :|||||
 DB 406 AspMetAsnProGlnLeuArgGluMetMetGlnAsnPro 418

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_rnp model

Run on: November 12, 2003, 06:19:14 ; Search time 46 Seconds
(without alignment)
8785.171 Million cell updates/sec

Title: US-09-735-251-3

Perfect score: 2302

Sequence: 1 GCGAGAGCGGAGAGAGAGAA...TGCCTTCATCTCCCTGTC 1271

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delcp 6.0 , Delcext 7.0

Searched: 1107863 seqs, 158726577 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_rnp_model -DEV=xlp
-C/cgn2.1/USPTO_speci_p/US09735251/runat_12112003_061634_28110/app_query_fasta_1.1415
-DB=A Geneseq_19Jun01 -QWTF=fastan -SUFFIX=rag -MINMATCH=0.1 -LOCPC=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -VATLX=blosum62 -TRANS=human45.cdi
-S=45 -DOCALIGN=200 -THR SCORE=100 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFN=pro -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09735251@cgn 1 1 81 -runat_12112003_061634_28110 -NCPU=6 -CPU=3
-NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPACK=100 -LNGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELCP=6 -DELCEXT=7

Database : A Geneseq_19Jun01

1: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1988.DAT
2: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1981.DAT
3: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1982.DAT
4: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1983.DAT
5: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1984.DAT
6: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1985.DAT
7: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1986.DAT
8: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1987.DAT
9: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1988.DAT
10: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1989.DAT
11: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1990.DAT
12: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1991.DAT
13: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1992.DAT
14: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1993.DAT
15: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1994.DAT
16: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1995.DAT
17: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1996.DAT
18: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1997.DAT
19: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1998.DAT
20: /SIDSI/cgdata/geneseq/geneseq_emb1/AA1999.DAT
21: /SIDSI/cgdata/geneseq/geneseq_emb1/AA2000.DAT
22: /SIDSI/cgdata/geneseq/geneseq_emb1/AA2001.DAT
23: /SIDSI/cgdata/geneseq/geneseq_emb1/AA2002.DAT
24: /SIDSI/cgdata/geneseq/geneseq_emb1/AA2003.DAT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	DB	ID	Description
1	2133	92.7	410	18	AAW15414	Activated platelet
2	2129	92.5	426	18	AAW15413	Human activated pl
3	1959	85.1	420	20	AAV42382	Amino acid sequenc
4	1959	85.1	420	22	AAJ39074	Human secreted pro
5	1959	85.1	420	23	ABW56783	Human polypeptide
6	1938	82.9	380	22	ABW56783	Human polypeptide
7	1686	73.2	341	23	ABW56783	Human protein Hpl0
8	94.1	40.9	191	22	AAW63948	Human polypeptide
9	757	32.9	164	22	AAW63948	Human prostate can
10	620	26.9	120	22	AAW63950	Human prostate can
11	610	26.5	128	22	AAW63951	Human polypeptide
12	483	20.5	110	22	AAW63951	Human prostate can
13	444	19.3	114	21	AAW58780	Breast and ovarian
14	389	16.9	76	21	AAW58780	Human secreted pro
15	172.5	7.5	547	22	ABW69125	Drosophila melanog
16	170.5	7.4	329	23	ABJ10942	Yeast selected int
17	164.5	7.1	536	21	AAW30040	Arabidopsis thalia
18	164.5	7.1	577	21	AAW30039	Arabidopsis thalia
19	162	7.0	693	23	ABW69529	Human polypeptide
20	161.5	7.0	1938	24	ABW6678	Streptomyces virid
21	159	6.9	783	19	AAW37151	Mouse neural Mena+
22	159	6.9	787	19	AAW37152	Mouse neural Mena+
23	159	6.9	802	19	AAW37153	Mouse neural Mena+
24	159	6.9	802	22	AAW37153	Mammalian enabled
25	158.5	6.9	1938	24	ABW6679	Streptomyces virid
26	157.5	6.8	245	22	ABG19161	Novel human diago
27	156.5	6.8	617	22	AAW01355	Human gene 4 encod
28	156.5	6.8	617	23	ABW79641	Human novel secret
29	156.5	6.8	617	23	ABW79641	Synoviolin amino a
30	156.5	6.8	617	23	ABW79641	Human albumin fusi
31	156	6.8	598	22	ABG14000	Novel human diago
32	152.5	6.6	441	22	ABW11413	Myobacterium tube
33	152.5	6.6	572	18	AAW31855	Human extensin hom
34	152.5	6.6	624	22	AAW47122	Human Chap1. Homo
35	152.5	6.6	624	23	ABW97373	Novel human protei
36	152.5	6.6	763	18	AAW31852	Myobacterium tube
37	152	6.4	1938	24	ABW6679	Streptomyces virid
38	152	6.4	1040	22	ABG14734	Novel human diago
39	150.5	6.5	819	22	ABG03569	Novel human diago
40	149	6.3	371	22	ABG03569	Novel human diago
41	149	6.5	415	22	ABW59727	Drosophila melanog
42	149	6.5	616	23	ABW83078	Synoviolin related
43	148	6.4	1023	23	AAW82954	Human homologue of
44	147.5	6.4	489	23	ABG4061	Human albumin fusi
45	147.5	6.4	505	22	AAW01380	Human gene 4 encod

ALIGNMENTS

RESULT 1
AAW15414
ID AAW15414 standard; Protein; 410 AA.
XX
AC AAW15414;
XX
DT 06-JUL-1997 (first entry)
XX
DE Activated platelet protein-2 APP-2 alternatively spliced variant.
XX
KW APP-2; activated platelet protein-2; thrombus; monoclonal antibody.
XX
OS Homo sapiens.
XX
PN W09712898-A1.
XX
PD 10-APR-1997.
XX

PF 04-OCT-1996; 96WO-US15922.
 XX
 PR 06-OCT-1995; 95US-0005074.
 XX
 PA (HARD) HARVARD COLLEGE.
 XX
 PI Clement Cy, Reed G;
 XX
 DR WPI; 1997-226153/20.
 DR N-PSDB; AAT64550.
 XX
 PT Naturally-occurring platelet activation polypeptide, APP-2 - used
 PT for locating a platelet thrombus in an animal, or activated platelet
 PT complex in a biological sample
 XX
 PS Claim 2; Fig 6; 50pp; English.
 XX
 CC A polypeptide (AAW15414) comprises an alternatively spliced variant
 CC of novel human activated platelet protein-2 (APP-2) (see also
 CC AAW15413), a protein that is preferentially expressed on activated
 CC human platelets but not resting platelets. Its amino acid sequence
 CC was deduced from an isolated cDNA clone (AAT64550). APP-2 is
 CC characterised as contg. an epitope that binds to monoclonal antibody
 CC 3B2 (ATCC CRU 11986). APP-2 can be produced in transformed host
 CC cells. Anti-APP-2 antibodies are useful in methods for detecting
 CC activated platelets in biological samples, for localising a platelet
 CC thrombus in an animal, and for detecting an activated platelet
 CC protein complex in an animal.
 XX
 SQ Sequence 410 AA;

Alignment Scores:
 Pred. No.: 6, 16e-177 Length: 410
 Score: 2133.00 Matches: 410
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.66% Indels: 0
 DB: 18 Gaps: 0

US-09-735-251-3 (1-1273) x AAW15414 (1-410)

QY 44 ATGCTCTCAGAGTGGGAGCTTGGTGAAGTGGTGGAGAGGAGGAGTCTTCTGCAAG 103
 Db 1 MetSerLeuSerAspThrPheSLeAlaValGlyLeuAlaAspGlnProLeuThrProLys 20
 QY 104 TCTATGCTCGGTGGGAGGAGGAGTGGTGAAGTGGTGGAGAGGAGTCTTCTGCAAG 163
 Db 21 SerIleLeuArgLeuProGluThrGlnGlnGlySerGlySerGlySerGlySerGly 40
 QY 164 TCATTCTGAGGAGCTTATGCTTGTGTAAGTGGTGGAGAGTGGTGGAGAGGAGTCTTCTGCAAG 223
 Db 41 SerPheLeuLysGlnLeuIleAlaGlyLysLeuSLeGlnSerValProAspProGluLeu 60
 QY 224 ATTGATCTGATCTACTGCTGGGAGGAGTGGTGAAGTGGTGGAGAGGAGTCTTCTGCAAG 283
 Db 61 IleAspLeuIleThrCysGlyArgGlyLeuLysAspGlnThrLeuAspPheThrGly 80
 QY 284 ATCAACCTGGTCCACTGCTGATGCTGCTGGAAGTGGTGGAGAGTGGTGGAGAGGAGTCTTCTGCAAG 343
 Db 81 IleGlnProGlySerThrValAlaMetArgGluPheArgValLeuHisThrAlaLeu 100
 QY 344 CCGAACCTGTGGACAAAGTGGTGGGAGTGGTGGAGAGTGGTGGAGAGGAGTCTTCTGCAAG 403
 Db 101 ProGluProValAspLysValAlaMetArgGluPheArgValLeuHisThrAlaLeu 120
 QY 404 CACAGAGTCTCTTACAGAGGAGGAGTGGTGGTGGAGAGTGGTGGAGAGGAGTCTTCTGCAAG 463
 Db 121 HisSerSerSerSerThrValAlaValPheLysMetLeuSerAsnLysGluSerLeu 140
 QY 464 GATCAGATCTTGTGGGAGGAGGAGGAGTGGTGGAGAGTGGTGGAGAGGAGTCTTCTGCAAG 523
 Db 141 AspGlnIleValAlaThrProGlyLeuSerSerAspProIleAlaLeuGlyValLeu 160

QY 524 CAGCAGAGGAGGAGTCTTCTGCTCTGCTGATCCCAATATGCTGATGATGCTGCTGCT 583
 Db 161 GlnAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspThrLeuValPro 180
 QY 584 GCTCACCAGGAGGAGTCTGCTCAATGCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 643
 Db 181 AlaHisProAlaLeuValAsnAlaIleValLeuValLeuHisSerValAlaGlySerAla 200
 QY 644 CCAATGCTGGGAGTCTGCTCTTCCCGAGGAGTCCCTCCAGCTCATACCCGGATATG 703
 Db 201 ProMetProGlyThrAspSerSerArgSerMetProSerSerSerSerSerSerSer 220
 QY 704 CCAGTGGCTCTCTGCTTGAAGGCTCTCAGATGATGAGGATGATGATGATGATGATGATG 763
 Db 221 ProGlyGlyPheLeuPheGluGlyLeuSerAspGluAspPheHisProAsnThr 240
 QY 764 AGTCTCACAGCTCTAGCAGTACTCCAGTCTCCGCTCCGCTCCGCTCCGCTCCGCTCCG 823
 Db 241 ArgSerThrProSerSerSerThrProSerSerArgProAlaSerLeuGlyThrSerGly 260
 QY 824 GCTGCTGGGAGGAGGAGTCTCAGCAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGG 883
 Db 261 AlaAlaGlyProArgProIleThrGlnSerGluLeuAlaThrAlaLeuAlaLeuAlaSer 280
 QY 884 ACTCCGAGAGGAGTCTCAGCAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGT 943
 Db 281 ThrProGlnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 300
 QY 944 TCACCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
 Db 301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAsnAspLeuPheSerGln 320
 QY 1004 GCCCTACAGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGT 1063
 Db 321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnThrPro 340
 QY 1064 CAGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGT 1123
 Db 341 GlnLeuGlnGlnLeuArgAspMetGlyIleGlnAspAspGluLeuSerLeuArgProCys 360
 QY 1124 AGGCAACGCTGGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAG 1183
 Db 361 ArgProProValGlyThrSerLysGlnProThrProSerSerSerSerSerSerSerSer 380
 QY 1184 GAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
 Db 381 GlnLeuProAlaSerProGluProProAlaSerCysArgGlyThrCysProThrProGluAla 400
 QY 1244 CTCATGAAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGTGGTGGAGGAGT 1273
 Db 401 LeuMetLysValProProSerLeuProVal 410

RESULT 2
 AAW15413
 ID AAW15413 standard; Protein; 426 AA.
 XX
 AC AAW15413;
 XX
 DT 06-JUL-1997 (first entry)
 XX
 DE Human activated platelet protein-2 APP-2.
 XX
 KW APP-2; activated platelet protein-2; thrombus; monoclonal antibody.
 XX
 OS Homo sapiens.
 XX
 PN W09712898-A1.
 XX
 PD 10-APR-1997.
 XX
 PF 04-OCT-1996; 96WO-US15922.
 XX
 PR 06-OCT-1995; 95US-0005074.

XX (HARD) HARVARD COLLEGE.

XX Clement CV, Read G;

XX WPI: 1997-226153/20.

XX N-PSDB: AAT64549.

XX Naturally-occurring platelet activation polypeptide, APP-2, used for locating a platelet thrombus in an animal, or activated platelet complex in a biological sample

XX Claim 1; Fig 5; 50pp; English.

XX A novel 25 kDa polypeptide (AAW15413) designated activated platelet protein-2 (APP-2) is preferentially expressed on activated human platelets but not resting platelets. APP-2 is characterized as contg. an epitope that binds to monoclonal antibody 3B2 (ATCC CRL 11986). Its amino acid sequence was deduced from a cDNA clone (AA64549) obtd. from a human bone marrow cDNA library. An alternatively spliced variant (see also AAW15414) was also isolated. CC APP-2 can be produced in transformed host cells. Anti-APP-2 CC antibodies are useful in methods for detecting activated platelets CC in biological samples, for localizing a platelet thrombus in an CC animal, and for detecting an activated platelet protein complex in CC an animal.

XX SQ Sequence 426 AA;

Alignment Scores:
Pred. No.: 1,39e-176 Length: 426
Score: 2129.00 Matches: 494
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 92.48% Indels: 0
DB: 18 Gaps: 0

US-09-735-251-3 (1-1273) x AAW15413 (1-426)

QY 44 ATGCTCTCTCAGACTGGCAGCTGGCGTGAAGTGGCTGACAGCCACTTACTCCAAAG 103
DB 1 MetSerLeuSerAspTrpHisLeuAlaValLysLeuAlaAspGlnProLeuThrProLys 20
QY 124 TCTATTTCTGGTGGCAGACACAAATGGGAATACTGCTAGGGGCTATAGTAT 163
DB 21 SerileLeuArgLeuProGlnThrProGlnGlyTyrSerLeuGlyTyrSerile 40
QY 164 TCATTTGTGAAGCACTATTCTGGGAGAAATGAGAGTCTTCCAGACCTGAGCTG 223
DB 41 SerPheLeuLysGlnLeulleAlaGlyLysLeuThrLeuSerAlaProAspProGluLeu 60
QY 224 ATTGATCTGATCTACTGTGGGAAGCTTAAAGATGACAGACACTTACTTATGCC 283
DB 61 IleAspLeulleTyrCysGlyArgLysLeuLysAspGlnThrLeuAspPheTyrGly 90
QY 284 ATTCAAGCTGGTCCACTGCTCATGTTCTGCGAAGTCTGGGCTGAACCTGATCAGAAA 343
DB 81 IleGlnProGlySerThrValHisValLeuArgLysSerTrpProGluProAspGlnLys 100
QY 344 CCGGAACCTGTGGCAAGTGGCTGCCATGAGAGTTCGGGTTGTCACACTGGCCCTG 403
DB 101 ProGluProValAspLysValAlaValArgGluPheArgValLeuHisThrAlaLeu 120
QY 404 CACAGAGCTCTCTTACAGGAGCGGCTTTTAAAGTGTCTAGCAATAAGAGTCTCTG 463
DB 121 HisSerSerSerTyrArgGluAlaValPheLysMetLeuSerAspLysGluSerLeu 140
QY 464 GATCAGATCATTTGGCCACCCAGGCTCAGAGTACCCCTATGCTCTTGGGTTCTC 523
DB 141 AspGlnIleValAlaThrProGlyLeuSerSerAspProIleAlaLeuGlyValLeu 160
QY 524 CAGGACAGGACCTCTCTCTGCTTCTGCTGATCCCAATATGCTTGATAGCTGGTGCCT 583

DB 161 GlnAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspThrLeuValPro 180
QY 584 GCTCACCAGCCCTCGTCAATGCCATTCTCTGGTTCCTGACTCTCGTAGCAGGAGTCCC 643
DB 181 AlaHisProAlaLeuValAsnAlaIleValLeuValLeuHisSerValAlaGlySerAla 200
QY 644 CCAATGCTGGGACTGACTCTCTTCCCGAGAGTCCCTCCAGCTCATACGGGATATG 703
DB 201 ProMetProGlyThrAspSerSerArgSerMetProSerSerSerTyrArgAspMet 220
QY 704 CCAGGTGGCTTCTCTTTGAAGGGCTCTCAGATGATGAGGATGCTTTTACCACCAACACC 763
DB 221 ProGlyGlyPheLeuPheGluGlyLeuSerAspGluAspPheHisProAsnThr 240
QY 764 AGGTCCACACCTCTAGCAGTACTCCAGCTCCCGCCAGCCCTCCCTGGGGTACAGTGA 823
DB 241 ArgSerThrProSerSerSerThrProSerArgProAlaSerLeuGlyTyrSerGly 260
QY 824 GCTGTGGCCCGCCGCCATCACCCAGAGTACAGTGGCCACCGCTTGGCCCTGGCCAGC 883
DB 261 AlaAlaGlyProArgProIleThrGlnSerGluLeuAlaThrAlaLeuAlaLeuAlaSer 280
QY 884 ACTCCGGAGAGCAGCTCTCACACACCGACTCTCTGGCACCCAGGGTCACTTCTCAGGGACC 943
DB 281 ThrProGluSerSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
QY 944 TCACCAATGCTCTCTGGTGTCCAGTACGGGAGCGCCATCACCAATGATCTTCTCAGCCAA 1003
DB 301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAsnAspLeuPheSerGln 320
QY 1004 GCCCTACAGCAGTCCCTTCAGGCTCTGGGAGCCAGCCCTTCAGAGCAGTGGCAGGCC 1063
DB 321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnTrpGlnPro 340
QY 1064 CAGCTGCCAGCAGTACGTACATGGGCATCCAGGACCATGAGTCCAGCTCGGCCCTGC 1123
DB 341 GlnLeuGlnGlnLeuArgAspMetGlyIleGlnAspGluLeuSerLeuArgProCys 360
QY 1124 AGCCACCGGTGGGACATCCAGCAGCCCTGGAGCTCATCTTTCTGGAGAGCCCAT 1183
DB 361 ArgProProValGlyThrSerLysGlnProTrpSerSerSerLeuGluGluProHis 380
QY 1184 GAATCCCTGCTTCCCTGACCCCGCCAGCCAGTTCACAGAGCTACTGCGCTTGGAGGCA 1243
DB 381 GluLeuProAlaSerProGluProProAlaSerCysArgGlyTyrCysProTrpGluAla 400
QY 1244 CTCATGAAGTGCCTCCATCTCTCCCT 1270
DB 401 LeuMetLysValProProSerLeuPro 409
RESULT 3
AAV42382
ID AAY42382 standard; Protein; 420 AA.
XX AC AAY42382;
XX DT 09-DEC-1999 (first entry)
XX DE Amino acid sequence of fx317_11.
XX secreted protein; cDNA library; clone; transmembrane protein;
KW signal sequence cloning; hybridization cloning; gene therapy;
KW receptor.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Key 229..241
XX Peptide /label= Leader/Signal peptide
XX Protein 242..420
XX /label= Mature protein
PN WO9942470-A1.

```

XX 26-AUG-1999.
XX 18-FEB-1999; 99WO-US03458
XX 18-FEB-1998; 98US-C075038
XX 17-FEB-1999; 99US-C251603
XX (GEMV) GENETICS INST INC.
XX Jacobs K, McCoy JM, LaValle EB, Collins-Racie DA, Nerberg D,
XX Treacy M, Agostino MJ, Steininger RC;
XX WPI; 1999-518580/43.
XX N-PSDB; AAZ20855.
XX New polynucleotides encoding human secreted proteins used for
XX therapeutic, diagnostic and research purposes.
XX Claim 14; Page 102; 125pp; English.
XX This is the amino acid sequence of the fx317_11 protein, which is
XX derived from the fx317_11 clone isolated from a human fetal brain cDNA
XX library.
XX The PNs and proteins of the invention are predicted to have biological
XX activities which would make them suitable for treating, preventing or
XX ameliorating medical conditions in humans and animals, although no
XX supporting data is given. Suggested activities include nutritional
XX activity, cytokine and cell proliferation/differentiation activity,
XX immune stimulating (e.g. as vaccines) or suppressing activity,
XX hematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumor invasion suppressor activity, and tumor
XX inhibition activity. The PNs are also stated to be useful for gene
XX therapy. Other activities include inhibiting the growth, infection or
XX function of bacteria, fungi, viruses and other parasites; effecting
XX bodily characteristics such as, e.g. weight, color, skin, etc.,
XX effecting biorhythms or cardiac cycles; enhancing fertility; treatment
XX of depression; treatment of pain, hormonal or endocrine activity.
XX Sequence 420 AA;

Alignment Scores:
Pred. No.: 8
Score: 1959.00
Percent Similarity: 93.24%
Best Local Similarity: 93.24%
Query Match: 95.10%
Indels: 2
Gaps: 0
DB: 20

US-09-735-251-3 (1-1273) x AAY42382 (1-420)
QY 17 AGNAGAGAGCGGAGAGAGAGAGATGTTCCTCAGATCGCATCGCGGTGAG 76
DB 32 Arg-Lys-Arg-Glu-Arg-Glu-Arg-Glu-Arg-Met-Ser-Leu-Ser-Arg-Phe-Ser-Leu-Ala-Val-Lys 51
QY 77 CTGGCTGACCGACCTTACTCCAAAGTCTATTTGGTTGTCAGACAGACAGAACTGGA 136
DB 52 Leu-Ala-Asp-Gln-Pro-Leu-Thr-Pro-Lys-Ser-Ile-Leu-Arg-Leu-Pro-Glu-Thr-Glu-Leu-Gly 71
QY 137 GAATCTCGTCTAGGGGTATAGTATTTCATTCTGAAAGCAGCTTATTCGTGGCAAACTC 196
DB 72 Glu-Tyr-Ser-Leu-Gly-Tyr-Ser-Ile-Ser-Phe-Leu-Lys-Gln-Leu-Ile-Ala-Gly-Lys-Leu 91
QY 197 CAGGAGTCTGTTCCAGACCTCAGCTGATGTCTGATCTACTGTGTCGAGCTAAAG 256
DB 92 Gln-Glu-Ser-Val-Pro-Asp-Pro-Glu-Leu-Ile-Asp-Leu-Ile-Tyr-Cys-Gly-Arg-Lys-Leu-Lys 111
QY 257 GATGACACACACTTGACTCTATCGATTCACCTGGTCCACTGTCTCATCTGTCGGA 316
DB 112 Asp-Asp-Gln-Thr-Leu-Asp-Phe-Tyr-Gly-Gln-Pro-Gly-Ser-Thr-Val-His-Val-Leu-Arg 131

```

```

QY 317 AAGTCTCGCTGAACTGATCAGAAACCGGAACTCTGTGGACAAAGTGGCTGCCATGAGA 376
DB 132 Lys-Ser-Tyr-Pro-Glu-Pro-Asp-Gln-Lys-Pro-Glu-Pro-Val-Asp-Lys-Val-Ala-Met-Arg 151
QY 377 GAGTCCCGGTGCTTGCACACTGCGCTGCACAGCAGCTCTCTTACAGGAGGCGGTCTTT 436
DB 152 Glu-Phe-Arg-Val-Leu-His-Thr-Ala-Leu-His-Ser-Ser-Ser-Tyr-Arg-Glu-Ala-Val-Phe 171
QY 437 AAGATGCTCAGCAATAAGGAGTCTCTGGATCAGATCATTTGTGGCCACCCCGGCTCAGC 496
DB 172 Lys-Met-Leu-Ser-Asn-Lys-Glu-Ser-Leu-Asp-Gln-Ile-Val-Ala-Thr-Pro-Gly-Leu-Ser 191
QY 497 AGTGACCTATTGCTCTTGGGTTCTCCAGAGCAAGACCTCTCTCTCTCTCTCTCTCTCT 556
DB 192 Ser-Asp-Pro-Ile-Ala-Leu-Gly-Val-Leu-Gln-Asp-Lys-Asp-Leu-Phe-Ser-Val-Phe-Ala-Asp 211
QY 557 CCCAATATCTGTATAGCTTGTGCTGCTCACCAGCCCTCGTCAATGCCAATGCTCTG 616
DB 212 Pro-Asn-Met-Leu-Asp-Thr-Leu-Val-Pro-Ala-His-Pro-Ala-Leu-Val-Ile-Ala-Leu-Val 231
QY 617 GTTCTGCACTCTGCTAGCAGGAGTGCCTCCCAATGCCCTGGGACTGACTCTCTCTCTCCCGGAGC 676
DB 232 Val-Leu-His-Ser-Val-Ala-Gly-Ser-Ala-Pro-Met-Pro-Gly-Thr-Asp-Ser-Ser-Arg-Ser 251
QY 677 ATGCCCTCAGCTCATACCGGATATGCCAGTGGCTTCTCTCTCTCTCTCTCTCTCTCT 736
DB 252 Met-Pro-Ser-Ser-Ser-Tyr-Arg-Asp-Met-Pro-Gly-Gly-Phe-Leu-Phe-Glu-Gly-Leu-Ser-Asp 271
QY 737 GATGAGGATGACTTTCACCCAAACACGAGTCCACACCTCTAGCAGTACTCTCCAGCTCC 796
DB 272 Asp-Glu-Asp-Phe-His-Pro-Ser-Thr-Arg-Ser-Thr-Pro-Ser-Ser-Ser-Thr-Pro-Ser-Ser 291
QY 797 CGCCAGCTCTCTCTGGGTACAGTGTGAGTGTGCGCCCGCCCGCCCATCACCCAGAGTCTGAG 856
DB 292 Arg-Pro-Ala-Ser-Leu-Gly-Tyr-Ser-Gly-Ala-Gly-Pro-Arg-Pro-Ile-Thr-Gln-Ser-Glu 311
QY 857 CTGGCCACCGCTTGGCCCTTGGCCAGCTCTCGGAGAGCAGCTCTCAGACACCGACTCTCT 916
DB 312 Leu-Ala-Thr-Ala-Leu-Ala-Leu-Ala-Ser-Thr-Pro-Glu-Ser-Ser-Ser-His-Thr-Pro-Thr-Pro 331
QY 917 GGCACACCGGCTACTCTCTCAGGAGCTCAGCAATGCTCTGCTGCTCCAGTCCAGGAGCG 976
DB 332 Gly-Thr-Gln-His-Ser-Ser-Gly-Thr-Ser-Pro-Met-Ser-Ser-Gly-Val-Gln-Ser-Gly-Thr 351
QY 977 CCATCAATATGATCTTTCAGCCGAGCCCTACAGCATGCCCTCTCAGGCTCTCTGCGGAG 1036
DB 352 Pro-Ile-Thr-Asn-Asp-Leu-Phe-Ser-Gln-Ala-Leu-Gln-His-Ala-Leu-Gln-Ala-Ser-Gly-Gln 371
QY 1037 CCGAGCTTCAGAGCCAGTGGGAGCCCGGAGCTGCAGCAGCTAGTGCACATGGGCATCCAG 1096
DB 372 Pro-Ser-Leu-Gln-Ser-Gln-Tyr-Gln-Pro-Gln-Leu-Gln-Gln-Leu-Arg-Asp-Met-Gly-Ile-Gln 391
QY 1097 GAGCATGAGCTGAGCTCGCGCC-CTGAGGCGCCAGCCGCTGGGAGCATCCAGCAGCCCTG 1155
DB 392 Asp-Asp-Glu-Leu-Ser-Leu-Arg-Ala-Leu-Gln-Ala-Thr-Gly-Gly-Asp-Ile-Gln-Ala-Leu 411
QY 1156 GAGCTCATCTTTCTCTGAGGAGCCCA 1182
DB 412 Glu-Leu-Ile-Phe-Ala-Gly-Gly-Ala-Pro 420

RESULT 4
AAU39074
ID AAU39074 standard; Protein; 420 AA.
XX AC AAU39074;
XX DT 16-JAN-2002 (first entry)
XX DE Human secreted protein fx317_11.
XX KW Human; secreted protein; antiinflammatory; immunosuppressive;
XX KW neutrotropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
XX KW cytostatic; antidiabetic; virucide; antinfertility; anticonvulsant;

```

vasotropic; antiparkinsonian; immunostimulant; dermatological;
antithematic; antitumor; antidiabetic; osteopathic; tranquilizer;
cerebroprotective; cytokine; cell proliferation; cell differentiation;
immune deficiency; severe combined immunodeficiency; SCID; tumor;
autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
Parkinson's disease; Huntington's disease; infection; cardiac disease;
stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
food supplement; vaccine.

OS Homo sapiens.

XX WO2001:75069-A2.

XX 11-OCT-2001.

XX 22-MAR-2001; 2001WO-US09369

XX 30-MAR-2000; 2000US-0539330

XX 04-DEC-2000; 2000US-0729674.

XX (GEVY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie E, Collins racie LA, Evans C,

PI Treacy M, Agostino MJ, Steininger RC, Spaulding V, Wong GO;

PI Clark H, Fichtel K, Verberg G;

XX WPI; 2001-639363/73.

XX N-PSDB; AAS5922.

PT Secreted human proteins, useful as vaccine for treating various
PT diseases such as autoimmune disorders (e.g. multiple sclerosis), and
PT nervous system disorders (e.g. stroke)

XX Disclosure: Page 578-579; 61pg; English.

CC The invention relates to novel human secreted proteins, the nucleic
CC acids encoding them. The protein may exhibit cytokine, cell proliferation
CC or cell differentiation activity or may induce production of other
CC cytokines in certain cell populations and may exhibit immune stimulating
CC or immune suppressing activity, which is useful for the treatment of
CC various immune deficiencies and disorders e.g. severe combined
CC immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis,
CC systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary
CC inflammation. The proteins are also useful in the treatment of diseases
CC and disorders including tissue, skin and bone regeneration and in
CC graft-versus-host diseases (GVHD) in the induction of tumor immunity,
CC myeloid or lymphoid cell deficiencies, wound healing and tissue repair,
CC in the treatment of burns, incisions and ulcers, as well as in treatment
CC of periodontal disease, osteoporosis or osteoarthritis, mediated by
CC inflammatory processes, diseases of the peripheral nervous system,
CC Alzheimer's, Parkinson's disease, Huntington's disease,
CC amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
CC infarction of cardiac and central nervous system vessels, e.g. stroke,
CC sepsis, inflammatory bowel disease, ulcers, bone regeneration. The
CC protein, having activin- or inhibin-related activities is useful as a
CC contraceptive based on the ability of inhibins to decrease fertility in
CC female mammals and decrease spermatogenesis in male mammals. The
CC proteins and nucleic acids are also useful as food supplements. The
CC present sequence represents a secreted protein of the invention.

XX Sequence 420 AA;

SQ

Alignment Scores:
Pred. No.: 8,71e-162 Length: 420
Score: 1959.00 Matches: 388
Percent Similarity: 99.74% Conservative: 0
Best Local Similarity: 99.74% Mismatches: 1
Query Match: 85.10% Indels: 1
DB: 22 Gaps: 0

US 09-735-251-3 (1-1273) x AAUS9074 (1-420)

QY 17 AGAAAGAGAGAGCCGAAAGAGAGAGAGATGCTCTCTCACACTGGCACCTGGCGGTGAAG 76
DB 32 Arglysa:qGluatGg.uA:Gg.uA:GmetSerLeuSerAspTrpHisLeuAlaValLys 51
QY 77 CTGGCTGACAGGCACCTACTCCAAACTCTATCTTCGGTTGCAGAGACAGAACTGGGA 136
DB 52 LeuAlaAspG.nPrLeuThrProLysSerileLeuArgLeuProGluThrGluLeuGly 71
QY 137 GAATACTGCTAGGGGGCTATAGTATTCTTCTGAAGCAGCTTATTGTCGCAACTC 196
DB 72 GluTyrSerLeuGlyG:YtyrSerileSerPheLeuLysGlnLeuileAlaGlyLysLeu 91
QY 197 CAGGAGTCTGTTCAGACCCCTGAGCTGATCTGATCTACTGTGTGTCGGAAGCTAAAA 256
DB 92 GlnGluSerValProAspProGluLeuileAspLeuileTyrCyaglyArglyLeuLys 111
QY 257 GATCACCAGACACTTACTTCTATGGCATTTCAACCTGGCTGGCTCCATCTCCATGTTCTCGA 316
DB 112 AspAspGlnThrLeuAspPheTyrGlyileGlnProGlySerThrValHisValLeuArg 131
QY 317 AAGTCTGGCCTGAACCTGATCAGAAACCGAACCTCTGCAACAAAGTGGCTGCATCAGA 376
DB 132 LysSerTrpProGluProAspGlnLysProGluProValAspLysValAlaAlaMetArg 151
QY 377 CAGTTCGGGTGTTCACACTGCCCTGCAGCAGCTCTCTTTACAGGAGGCGGTCTTT 436
DB 152 GluPheArgValLeuH:sthrAlaLeuHisSerSerSerTyrArgGluAlaValPhe 171
QY 437 AAGATGCTCAGCAATAAGGAGTCTCTCGATCAGATCATTGTGGCCACCCAGCCCTCAGC 496
DB 172 LysMetLeuSerAsnLysGluSerLeuAspGlnlelleValAlaThrProGlyLeuSer 191
QY 497 AGTCACCTTATGCTCTTGGGGTTCTCCAGGACAAAGACCTCTTCTCTGCTCGCTGAT 556
DB 192 SerAspProileAlaLeuGlyValLeuGluAspLysAspLeuPheSerValPheAlaAsp 211
QY 557 CCCAATA:GCTTGATAGTGTGGTGCCTGCTCACCAGCCCTCTGTCATGCAATGCTGCTG 616
DB 212 ProAsrMetLeuAspThrLeuValProAlaHisProAlaLeuValAsnAlaileValLeu 231
QY 617 GTTCTGCACTCCGTAGCAGGAGTGGCCCCAATGCTGGGACTGACTCTCTTCTCCCGGAGC 676
DB 232 ValLeuH:ssrValAlaGlySerAlaProMetProGlyThrAspSerSerArgSer 251
QY 677 ATGCCCTCACTCATACCGGATATCCAGGTGGCTCTCTGTTGAAGGGCTCTCAGAT 736
DB 252 MetProSerSerSerTyrArgAspMetProGlyglyPheLeuPheGluGlyLeuSerAsp 271
QY 737 GATGGGATGACTTTCACCCCAACACAGGTCGACACCTCTAGCAGTACTCCAGGCTCC 796
DB 272 AspGluAspAspPheHisProAsnThrArgSerThrProSerSerSerThrProSerSer 291
QY 797 CCCCAGGCTCTCTGGGGTACAGTGGAGTGTGGGCCCCGGCCCATCACCAGAGTGAG 856
DB 292 ArgProAlaSerLeuGlyTyrSerGlyValAlaGlyProArgProIleThrGlnSerGlu 311
QY 857 CTGGCCACCGCTTGGCCCTGGCCGAGCAGCTCCCGAGAGAGAGCTCTCACACCGACTCTCT 916
DB 312 LeuAlaThrAlaLeuAlaLeuAlaSerThrProGluSerSerSerHisThrProThrPro 331
QY 917 GGCACCCAGGCTCATTCTCAGGAGCCTCACCAGTCTCTGGTGTCCAGTCCAGGGAGC 976
DB 332 GlyThrGlnGlyHisSerSerGlyThrSerProMetSerSerGlyValGlnSerGlyThr 351
QY 977 CCCATCACCAATGATCTCTTTCAGGCAAGCCCTACAGCATGCCCTTACGGCTCTGGGCGAG 1036
DB 352 ProIleThrAsnAspLeuPheSerGlnAlaLeuGlnHisAlaLeuGlnAlaSerGlyGln 371
QY 1037 CCCAGCCTTCAGAGCCAGTGGCAGCCCGCCAGCTGAGCAGCTCTGAGTGGGATCCAG 1096
DB 372 ProSerLeuGlnSerGlnTrpGlnProGlnLeuGlnGlnLeuArgAspMetGlyileGln 391

QY 1097 GACGATGAGCTGGGCG-CTGAGGAGGACGGTGGGACATCCAGGACCGCTG 1155
 DB 392 AspAspGluLeuSerLeuArgAlaLeuGluA:ThrG:YGIYAspIleGleA:AlaLeu 411
 QY 1156 GAGCTCATCTTTGCTGGAGAGCCCA 1182
 DB 412 GluLeuIlePheAlaGlyAlaPro 420
 RESULT 5
 ID ABB55783
 AC ABB55783 standard; Protein: 420 AA.
 XT ABB55783:
 DT 14-FEB-2002 (first entry)
 DE Human polypeptide SEQ ID NO 172.
 XX Human: clone bd306-7; clone yb8-1; ATCC number 98599; gene therapy;
 KW immune disorder; bacterial infection; fungal infection; cancer; tumour;
 KW autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
 KW osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
 KW haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulnery;
 KW ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
 KW Crohn's disease; cytostatic; anti-inflammatory; immunomodulator;
 KW neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
 XX Homo sapiens.
 OS
 XX US2501039335-A1.
 XX 08-NOV 2001.
 XX 04 DEC-2000, 2000US-0729674.
 XX 10-APR-1997; 97US-126425P.
 PR 04-DEC-1997; 97US-067454P.
 PR 20-DEC-1997; 97US-069379P.
 PR 02-JAN-1998; 98US-070346P.
 PR 07-JAN-1998; 98US-070643P.
 PR 08-JAN-1998; 98US-070755P.
 PR 12-JAN-1998; 98US-071134P.
 PR 13-JAN-1998; 98US-071334P.
 PR 30-JAN-1998; 98US-071395P.
 PR 18-FEB-1998; 98US-075014P.
 PR 30-MAR-2000; 2000US-053911.
 PR 23-NOV-1998; 98US-079758P.
 XX JACO/J; JACOBS K.
 PA (MCO/J) MCCOY J M.
 PA (LAVA/J) LAVALLIE E R.
 PA (COLL/J) COLLINS-RACIE L A.
 PA (EVAN/J) EVANS C.
 PA (MERB/J) MERBERG D.
 PA (TREA/J) TREACY M.
 PA (AGOS/J) AGOSTINO M C.
 PA (STEI/J) STEININGER R J.
 PA (SPAU/J) SPAULDING V.
 PA (WONG/J) WONG G G.
 PA (CLAR/J) CLARK H.
 PA (FECH/J) FECHTEL K.
 XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
 PI Wong GG, Clark H, Fechtel K;
 XX WPI: 2002-040725/05.
 DR N-PSDB; ABA90961.
 XX New secreted proteins and encoding polynucleotides, useful in gene
 PT therapies, particularly for preventing or treating autoimmune
 PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,

stroke or inflammations
 Disclosure: Page 309-310; 349pp; English.
 The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
 ABA90980) and encoded proteins (ABB55698-ABB55800), especially
 polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
 proteins SEQ ID NO 2 (ABB55698) and SEQ ID NO 20 (ABB55707) contained in
 clones bd306-7 and yb8-1 respectively and the clones bd306-7 and yb8-1
 are deposited with the American Type Culture Collection (ATCC) with
 accession number 98599. The polynucleotides and encoded polypeptides have
 cytostatic, anti-inflammatory, immunomodulator, vulnery,
 neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
 and anti-inflammatory activity and acting as cytokine modulators,
 haematopoiesis regulators, tissue growth modulators and/or cadherin
 suppressors. The polypeptides and polynucleotides are useful in gene
 therapies, particularly for preventing, treating or ameliorating any of
 the following diseases: immune deficiency and disorders; e.g. bacterial
 or fungal infections, autoimmune disorders, cancer, systemic lupus
 erythematosus or graft-versus-host disease, myeloid or lymphoid cell
 deficiencies; wound, burns, incisions and ulcers, osteoporosis or
 osteoarthritis; central and peripheral nervous system diseases and
 neuropathies; e.g. Alzheimer's, Parkinson's disease, Huntington's
 disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
 haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
 or systemic inflammatory response syndrome; ischaemia-reperfusion
 injury, endotoxin lethality, arthritis, inflammatory bowel disease or
 Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
 foliaceus.
 XX
 SQ Sequence 420 AA;
 Alignment Scores:
 Pred. No.: 8,71e-162 Length: 420
 Score: 1959.03 Matches: 388
 Percent Similarity: 99.74% Conservative: 0
 Best Local Similarity: 99.74% Mismatches: 1
 Query Match: 85.10% Indels: 1
 DB: 23 Gaps: 0
 US-09-735-251-3 (1-273) x ABB55783 (1-420)
 QY 17 AGAAGAGAGAGCGCAAGAGAGAGAGATGCTCTCTCAGACTGGCAGCTGGCGTGAAG 76
 DB 32 ArglyAspAlaGluArgGluArgGluArgGluSerLeuSerAspTrpHisLeuAlaValLys 51
 QY 37 CTGGATCAGAGAGCACTTACTCTCAAGAGTCTATTCTTCGGTTCAGAGAGAGAGAGCTGGGA 136
 DB 52 LeuAlaAspGlnProLeuThrProLysSerIleLeuArgLeuProGluThrGluLeuGly 71
 QY 137 GAATCTCTGGTAGGGCTATAGTATTTTCATTTCTCAAGCAGCTTATTCTGCGCAACTC 196
 DB 72 GluTyrSerLeuGlyGlyTyrSerIleSerPheLeuLysGlnLeuIleAlaGlyLysLeu 91
 QY 197 CAGGAGTCTTCTCAGACCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 256
 DB 92 GluGluSerValProAspProGluLeuLeuAspLeuIleTyrCysGlyArgLysLeuLys 111
 QY 257 GATGACCCAGACACTTCACTTCTATGCGATTCATCACTGGGTTCCTCACTGTCCTGCGCA 316
 DB 112 AspAspGlnThrLeuAspPheTyrGlyIleGlnProGlySerThrValHisValLeuArg 131
 QY 317 AAGTCTCTGGCTCAACTGATCAGAAACCGGAACTGTGCAAACTGCTGCTGCTGCTGCTGCTGCT 376
 DB 132 LysSerTrpProGluProAspGlnLysProGluProValAspLysValAlaAlaMetArg 151
 QY 377 GAGTTCGGGTGTTCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 436
 DB 152 GluPheArgValLeuHisThrAlaLeuHisSerSerSerSerSerSerSerSerSerSerSer 171
 QY 437 AAGATCTCAGCAATAGGAGTCTCTGGATCAGATCATTTGCGCCACCCAGGCTCAGC 496
 DB 172 LysMetLeuSerAsnLysGluSerLeuAspGlnIleValAlaThrProGlyLeuSer 191

QY 497 AGTGACCTATTGCTTGGGGTCTCCAGACAGGACCTCTTCTCTGCTTGGCTGAT 556
 DB 192 SerAspProIleAlaLeuGlyValLeuGlnAspLysAspLeuPheSerValPheAlaAsp 211
 QY 557 CCCATATGCTTGATACCTGGTGGCTGCTCACCAGCCCTCGTCAATGCGATGCTG 616
 DB 212 ProAsnMetLeuAspThrLeuValProAlaHisProAlaLeuValAsnAlaIleValLeu 231
 QY 617 GTTCTGCACTCCGTAGCAGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676
 DB 232 ValLeuHisSerValAlaGlySerAlaProMetProGlyThrAspSerSerSerArgSer 251
 QY 677 ATGCCCTCAGCTCATACCGGATATGCCAGTGGCTTCCTGTTTGAAGGCTCTCAGAT 736
 DB 252 MetProSerSerSerTyArgAspMetProGlyGlyPheLeuPheGlyLeuSerAsp 271
 QY 737 GATGAGGATGCTTTCACCCAAACACACAGGTCCACACCTCTACGACTACTCCAGCTCC 796
 DB 272 AspGluAspAspPheHisProAsnThrArgSerThrProSerSerSerThrProSerSer 291
 QY 797 CGCCAGCTCCCTGGGTACAGTGGAGCTCTGGGCTGGGCTGGGCTGGGCTGGGCTG 856
 DB 292 ArgProAlaSerLeuGlyTySerGlyAlaAlaGlyProArgProIleThrGlnSerGln 311
 QY 857 CTGGCCACCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 916
 DB 312 LeuAlaThrAlaLeuAlaLeuAlaSerThrProGluSerSerSerHisThrProThrPro 331
 QY 917 GGCACCCAGGCTCATCTCCAGGACCTCCACCAATGCTCTGCTGGTGTCCAGTCAGGACG 976
 DB 332 GlyThrGlnGlyHisSerSerGlyThrSerProMetSerSerGlyValGlnSerGlyThr 351
 QY 977 CCCATCAACCAATATGCTTTCAGGCAAGCCCTACAGATGCCCTTCAGGCTCTGGGACG 1036
 DB 352 ProIleThrAsnAspLeuPheSerGlnAlaLeuGlnHisAlaLeuGlnAlaSerGlyGln 371
 QY 1037 CCCAGCTTCAGACCTGAGCCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1096
 DB 372 ProSerLeuGlnSerGlnThrProGlnLeuGlnLeuAlaGlnMetGlyIleGln 391
 QY 1097 GACGATGAGCTGAGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGGCTCCGG 1155
 DB 392 AspAspGluLeuSerLeuAlaGlnAlaLeuGlnAlaThrGlyGlyAspIleGlnAlaAlaLeu 411
 QY 1156 GAGCTCATCTTGTGGAGGAGCTGCA 1182
 DB 412 GluLeuIlePheAlaGlyAlaAlaPro 426

RESULT 6
 AAG93292
 ID AAG93292 standard; Protein: 380 AA.
 XX AC AAG93292;
 AC AC AAG93292;
 DT 13-SEP-2001 (first entry)
 XX DE Human protein HP:0657.
 XX KW Human; gene therapy; tumour.
 XX OS Homo sapiens.
 XX PN WO2001:42302-A1.
 XX PD 14-JUN-2001.
 XX PF 06-DEC-2000; 2000WO-JP08631.
 XX PR 06-DEC-1999; 99JP-0346863.
 PR 06-DEC-1999; 99JP-0346864.
 PR 08-FEB-2000; 2000JP-0031062.
 PR 10-FEB-2000; 2000JP-0034090.

PR 10-FEB-2000; 2000JP-0034091.
 PR 14-FEB-2000; 2000JP-0035829.
 PR 14-FEB-2000; 2000JP-0035899.
 PR 14-MAR-2000; 2000JP-0071161.
 PR 30-MAY-2000; 2000JP-0160851.
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PA Kato S, Eguchi C, Saeki M;
 PI WPI: 2001-381646/40.
 DR N-PSDB; AAH68577.
 XX Human protein originated from tumor cell line, applicable as drug,
 PT reagent for studying intracellular protein networks and protein source
 PT for drug screening, also encoded cDNA for gene diagnosis and gene
 PT therapy
 XX Claim 1: Pages 313-314; 471pp; Japanese.
 CC The present sequence is a human protein. The human protein, preferably
 CC originated from tumor cell line, is applicable as a drug, a reagent for
 CC studying intracellular protein networks and a protein source for
 CC screening proteins for binding low molecular weight drugs. The human
 CC protein coding sequence is useful for gene diagnosis and gene therapy,
 CC expression vectors and transformant cells for detection of ligands and
 CC receptors.
 XX SQ Sequence 380 AA;
 Alignment Scores:
 Pred. No.: 2,31e-157 Length: 380
 Score: 1908.00 Matches: 378
 Percent Similarity: 99.47% Conservative: 0
 Best Local Similarity: 99.47% Mismatches: 2
 Query Match: 82.88% Indels: 1
 DB: 22 Gaps: 0
 US-09-735-251-3 (1-1273) x AAG93292 (1-380)
 QY 44 ATGTCTCTCTCAGACTGGCACCCTGGCGGTGAAGCTGGCTGACCCAGCCACTTACTCCAAG 103
 DB 1 MetSerLeuSerAspTrpHisLeuAlaValLysAlaAspGlnProLeuThrProLys 20
 QY 104 TCTATTCTTCGTTCCGACAGACAGAACTGGAGAAATACTCCTAGGGGGCTATAGTATT 163
 DB 2 SerIleLeuArgLeuProGluThrGluLeuGlyGlySerLeuGlyGlyTySerIle 40
 QY 164 TCATTTCGAAACAGCTTATTGCTGGCAAACTCCAGGAGTCTGTTCACAGACCTGAGCTG 223
 DB 41 SerPheLeuLysGlnLeuIleAlaGlyLysLeuGlnGluSerValProAspProGluLeu 60
 QY 224 ATTGATCTGATCTACTGTGGTGGAGAGCTAAAGATGACACAGACACTTCTATGCTG 283
 DB 61 IleAspLeuIleTyCysGlyArgLysLeuLysAspAspGlnThrLeuAspPheTyGly 80
 QY 284 ATTCACCTGGGTCCACTGCTCATGTTCTGCGAAAGTCTGCTGCTGCTGCTGCTGCTG 343
 DB 81 IleGlnProGlySerThrValHisValLeuArgLysSerTrpProGluProAspGlnLys 100
 QY 344 CCGGAACCTGTGGCAAAAGTGGCTGCCATGAGAGAGTTCGGGTGTTTCCACACTGCCCTG 403
 DB 101 ProGluProValAspLysValAlaAlaMetArgGluPheArgValLeuHisThrAlaLeu 120
 QY 404 CACAGAGCTCTCTTACAGGAGCGCTCTTTAAGATGCTCAGCAATTAAGAGTCTCTG 463
 DB 121 HisSerSerSerTyArgGluAlaValPheLysMetLeuSerAsnLysGluSerLeu 140
 QY 464 GATCAGATCATGTGGCCACCCAGGCTTCAGAGTGGCTTCTTCTCTTCTGCTGCTGCT 523
 DB 141 AspGlnIleIleValAlaThrProGlyLeuSerSerAspProIleAlaLeuGlyValLeu 160
 QY 524 CAGGACAGGACCTCTCTCTGCTCTCTGATCCCAATAAGCTTGTGATACCTGCTGCT 583

DR WP1: 2001-025274/03.
 XX Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 XX cancer.
 XX
 XX Example 1: Page 790; 799pp; English.
 XX
 XX AAF22422 to AAF22626, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleotide sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAPs) respectively.
 CC AAF23015 to AAF23467, AAF23468 to AAF23701 and AAF23702 to AAF23970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytosolic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterised by expression of an abnormal amount of a protein,
 CC e.g. cancer.
 XX
 XX Sequence 110 AA;
 SQ
 Alignment Scores:
 Pred. No.: 1,59e-33 Length: 110
 Score: 483.00 Matches: 91
 Percent Similarity: 88.99% Conservative: 4
 Best Local Similarity: 85.32% Mismatches: 11
 Query Match: 20.49% Indels: 2
 DB: 22 Gaps: 0
 US-09-735-251-3 (1:1273) x AAF23951 (1:110)
 QY 934 GAATGACCTGGTGGCCAGAGAGGGTGTGTGAGAGTGTGTTCGGAGTGTGGCCAG 875
 Db 1 GlyMetThrLeuGlyAlaArgSerArgCysValArgAlaLeuArgSerAlaGlyCin 20
 QY 974 GGCAGGCGGTGGCCAGTCACTTGGGTATGGGCGAGTCCAGAGTCCACTGTA 815
 Db 21 GlyGlnGlyGlyGlnLeuThrLeuGlyAlaGlyProGlyPtoSerSerThrVal 40
 QY 814 CCCCAGGAGGTGGCGGAGTCCAGTATGCTATGCTATGCTATGCTATGCTATGCT 756
 Db 41 ProGlnGlyGlyTrpAlaGlyAlaGlySerThrAlaArgValTrpThrTrpCysLeu 60
 QY 755 GGTGAAGTCACTCCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 696
 Db 61 GlyGluSerHisProHisHisLeuAlaThrLeuGlnThrGlySerPheTrpHisLeuP 80
 QY 695 GGTATGAGTGGAGGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 636
 Db 80 oValLeuAlaGlyHisAlaHisAlaHisAlaHisAlaHisAlaHisAlaHisAla 100
 QY 635 CTGCTACGAGTGCAGACACGAGAC 611
 Db 100 acysPheArgValAsnGlnGlu 108
 RESULT 13
 AAB58780
 ID AAB58780 standard; Protein: 114 AA.
 XX
 XX AAB58780;
 XX
 XX 27-MAR-2001 (first entry)
 XX
 XX Breast and ovarian cancer associated antigen protein sequence SEQ ID 488.
 XX
 XX Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neotropic; neuroprotective; antiviral; antiparasitic and cardiac activity. The
 KW antidiabetic; antinflammatory; antitumor; antitoxic; anticonvulsant;
 KW antidiabetic; antinflammatory; antitumor; antitoxic; anticonvulsant;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia.
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis.

KW cardiovascular disorder; wound healing; neurological disease.
 XX
 XX Homo sapiens.
 XX WO2000055173-A1.
 PN
 XX 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2500WC-US05581.
 XX
 XX 12-MAR-1999; 99US-0124270.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Rosen CA, Ruben SM;
 XX WPI: 2000-611515/58.
 DR N-PSDB; AAF21683.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases.
 XX
 XX Claim 11: Page 919; 1299pp; English.
 XX
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neotropic; neuroprotective; antiviral; antiparasitic; anticonvulsant;
 CC antidiabetic; antinflammatory; antitumor; antitoxic; anticonvulsant;
 CC antidiabetic; antinflammatory; antitumor; antitoxic; anticonvulsant;
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 XX Sequence 114 AA;
 SQ
 Alignment Scores:
 Pred. No.: 3,99e-30 Length: 114
 Score: 444.00 Matches: 92
 Percent Similarity: 88.79% Conservative: 3
 Best Local Similarity: 85.98% Mismatches: 10
 Query Match: 19.29% Indels: 2
 DB: 21 Gaps: 0
 US-09-735-251-3 (1:1273) x AAB58780 (1:114)
 QY 8 GGCAGAGGAG 67
 Db 10 GlyThrGlyArgGluArgGluArgGluArgGluArgGluArgGluArgGluArgGlu 29
 QY 68 GCGGTGAAGCTGGCTGACAGGACCTTACTCCAAAGTCTATCTTCGTTGCCAGACAC 127
 Db 30 AlaValLysLeuAlaAspGlnProLeuAlaProLysSerLeuGlnLeuProGluSer 49
 QY 128 GAATGGAGAGTACTCGCTAGGGGCTATAGTATTTCTTTCAGCAGCTTATGCT 187
 Db 50 GluLeuGlyGlyTrpSerLeuGlyGlyTrpSerLeuGlyGlyTrpSerLeuGlyGly 69
 QY 188 GGCAGAGTCCAGGAGTCTGTTCAGACCCCTGAGCTCTATTCATCTACTGCTGCTG 247
 Db 70 GlyLysLeuGlnGluSerValProAspProGluLeuLeuLeuLeuLeuLeuLeuLeu 99

QY 248 AAGCTAAAGATGACACAGACTTCTATGCACTTCAAGCTGGCTCACTGTCCAT 307
 Db 90 LysLeuLysAspAsp...Thr-LeuThrSerThrVal-PheAsnLeuAlaPro-HisProC 109
 QY 308 GTTCTGCGAAAGCTCTG 324
 Db 109 ysSer...Glu...Leu 114

RESULT 14

AAG02759
 ID AAG02759 standard; Protein: 76 AA.
 XX
 AC AAG02759;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein, SEQ ID NO: 6840.
 XX
 KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.
 XX
 OS Homo sapiens.
 XX
 PN EP1033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0203610.
 XX
 PR 26-FEB-1999; 99US-0122487.
 XX
 PA (GEST) GENSET.
 XX
 PI Dumas Milne Edwards J, Ducloux A, Giordano J;
 XX
 DR WPI: 2000-500381/45.
 DR N-PSDB; AAC02765.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 6840; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or poly(A) RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (3'UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC 3'UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 76 AA;

Alignment Scores:

Pred. No.:	2,098-25	Length:	76
Score:	389.00	Matches:	76
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	16.90%	Indels:	0
DB:	21	Gaps:	0

US-09-735-251-3 (1-1273) x AAG02759 (1-76)

QY 44 ATGTCTCTCTCAGACTGGCAGCTGCGGTGAAGCTGGCTGACACGCACTTACTCCAAAG 103
 |||||

Db 1 MetSerLeuSerAspTrpHisLeuAlaValLysLeuAlaLeuGlnProLeuThrProLys 20
 QY 104 TCTATCTCTCGTTCGCCAGACAGACTGGAGAATCTCGTAGGGGCTATAGTATT 163
 Db 21 SerIleLeuArgLeuProGluThrGluLeuGlyGlySerLeuGlyGlyTyrSerIle 40
 QY 164 TCATTTCGAAAGCAGCTTATTGCTGGCAAACTCCAGGAGTCTGTTCAGACCCCTGAGCTG 223
 Db 41 SerPheLeuLysGlnLeuIleAlaGlyLysLeuGlnGluSerValProAspProGluLeu 60
 QY 224 ATTGATCTGATCTACTGTGTCGGAAGCTAAAGATGACCCAGACACTT 271
 Db 61 IleAspLeuIleTyrCysGlyArgLysLeuLysAspAspGlnThrLeu 76

RESULT 15

ABB69125
 ID ABB69125 standard; Protein: 547 AA.

XX
 AC ABB69125;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 34167.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW Pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 21-MAR-2001; 2001WC-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL13228.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 34167; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 547 AA;

Alignment Scores:

Pred. No.:	3,16e-06	Length:	547
Score:	172.50	Matches:	109
Percent Similarity:	35.30%	Conservative:	82
Best Local Similarity:	20.15%	Mismatches:	153
Query Match:	7.49%	Indels:	198
DB:	22	Gaps:	23

US-09-735-251-3 (1-1273) x ABB69125 (1 547)

```

QY 95 ACTCCAAAGTCTATTCTTCGGTTCCACAGACAGAACTGGGAGATATCTCGTAGGGGC 154
DB |||||
DB 15 ThrProLysAsp-----LysLysThrValGluValAspGluAspSer 28
QY 155 TATATATTTCATTCTTAAGCACTTATTGTGCAAACTGACAGAGTGTGTTCAGAC 214
DB |||||
DB 29 --GlyIleLysAspPheLysIleLeuValAlaGlnLysPheGlu --AlaGlu 44
QY 215 CTGAGCTGATTGATCTGATCTA TGTGCTGGAGAGCTAAAGATGACAGACACTTGAC 274
DB |||||
DB 45 ProGluGlnLeuValLeuIlePheAlaGlyLysIleMetLysAspThrAspThrLeuGln 64
QY 275 TTCTATGGCATTCAACCTGGTCCACTGTCCTGATGCTGCAAGTCTGCTGCAACT 334
DB |||||
DB 65 MetHisAsnIleLysAspAsnLeuThrValHisLeuValIleLysAlaProThrArgAsn 84
QY 335 GATCAGAAACCG-----CAACCTGTGGAC----- 358
DB |||||
DB 85 AsnGluGlnProAlaArgGlnProAlaAspValArgGlnThrProPheIleLeuAsnGln 104
QY 359 -----AAGTGGCTGCCATGAGA----- 376
DB |||||
DB 105 PheGlyGlyLeuAlaGlyMetGlnAlaLeuGlyAlaGlySerAsnThrPheMetAspLeu 124
QY 377 GAGTTCGGGGTGTTCACACTGCTGTGCACACAGCACTCTCTTACAGGAG----- 427
DB |||||
DB 125 GlnAlaArgMetGlnAsnGlnLeuLeuAsnAsnGlyAspValLeuArgSerLeuMetAsp 144
QY 428 -----GGGCTTTTAAATGCTAGCAATAAGGAGTCTCTGATTCAGATCATTTGGCC 481
DB |||||
DB 145 AsnProMetValGlnGlnMetMetAsnAsnProAspThrValArgGlnLeuIleThrSer 164
QY 482 ACCCAGGC-----CTAGCACTGCACTATGTTCTGGGGTCTCCAGGAC 529
DB |||||
DB 165 AsnProGlnMetHisAspLeuMetGlnArgAsnProGluIleSerHisMetLeuAsnAsn 184
QY 530 AAGGACCTC-----TTCTGTCTTGGCTGATGCTCAATGCTGATACGTTG 577
DB |||||
DB 185 ProAspLeuLeuArgGlnThrMetGluLeuAlaArgAsnProSerValLeuGlnGluLeu 204
QY 578 GTGCTGTCTAC-----CCAGCCCTCGTCAATGCC 607
DB |||||
DB 205 MetArgSerHisAspArgAlaMetGlnAsnGlnLeuSerValProGlyLysSerAla 224
QY 608 ATTGCTCTGCTCTGCACTG----- 631
DB |||||
DB 225 LeuGlnArgIleThrArgAspThrIleGlnMetValAlaAlaAlaThrGlnSerPhe 244
QY 632 -----GCAATGAGTCCGCA 646
DB |||||
DB 245 GlyArgAsnProPheAlaGlyLeuValAspGlyGlyGlySerGlyValAlaGlyAsnAsnPro 264
QY 647 ATGCTGTGGAGTACTCTCTTCCCGGAGATGCCC-----TCCAGC 688
DB |||||
DB 265 GlnGlnGlyThrGluAsnArgAsn---ProLeuProAsnProIlePheGlyValAlaAsnSer 283
QY 689 TCATACCGGATATGCCAGTGGATCTCTGTTGAGGGCTTCACATGATGAGGATGAC 748
DB |||||
DB 284 GlyThrAsnGlyThrValGlyGly-----SerGlyValAlaGlyAsnProThrGlyAsp 300
QY 749 TTTCACCCA-----AACACAGGTCCACACCTCT----- 778
DB |||||
DB 301 LeuProProAsnAsnValLeuAsnThrProAlaMetArgSerMetLeuGlnMetAla 320
QY 778 ----- 778
DB |||||
DB 321 AspAsnProAlaMetMetGlnAsnLeuLeuAsnAlaProThrThrArgSerMetMetGlu 340
QY 779 -----AGCAGTACTCCC-----AGC 793
DB |||||
DB 341 SerMetSerGlnAspProAspMetAlaAlaArgLeuLeuSerSerProLeuMetSer 360

```

```

QY 794 TCCGCCCAAGCTCCCTCGGGTACAGTGGAGCTGCTGGGCCCGCCCATCACCAGAGT 853
DB |||||
DB 361 AsnAsnProAlaLeuGlnGluGlnValArgGlnMetMetProGlnPheMetAlaGlnMet 380
QY 854 GAGTGGCCACCCCTTGGCCCTGGCCAGCAGCTCCGGAGAGCAGCTCT----- 901
DB |||||
DB 381 GlnAsnProGluValMetAsnMetLeuThrAsnProAspAlaMetAsnAlaIleLeuGln 400
QY 902 -----CACACACCACTCTCTGGC----- 919
DB |||||
DB 401 IleGlnGlnGlyMetGluGlnLeuArgSerAlaAlaProGlyLeuValGlyThrLeuGly 420
QY 920 -----ACCCAGGCTCATCTCTCAGGACCTCCACCAATG----- 952
DB |||||
DB 421 IleProProProProGlyAlaGlyThrGlyThrAsnProAlaSerGlyAspGlySer 440
QY 953 -----TCTCTGCTGTCTCCAGTCAGG 973
DB |||||
DB 441 GlyIleAsnSerGlyAlaSerThrAsnAsnValSerProSerSerGlyLeuAsnAlaGly 460
QY 974 ACGCCCATCACCAT-----CATCTCTCAGCCAGGCC 1006
DB |||||
DB 461 ThrGlyThrProAsnLeuAlaProGlyGlyGlyProAsnAlaGlnLeuPheAsnAspPhe 480
QY 1007 CTACAGCATGCTCTCAGGCC-----TCTGGCAGCCCGCCAGCTTCAG 1048
DB |||||
DB 481 MetMetArgMetLeuAsnGlyMetSerAsnAsnAlaAspAsnThrGlnProPro---Glu 499
QY 1049 AGCCAGTGGCAGCCCGCAGCTGCGAGCAGCTACGTGACATGCGGCATCCAGCAAGATGAGCTG 1108
DB |||||
DB 500 ValArgTyrGlnSerGlnLeuGluGlnLeuAsnAlaMetGlyPheValAsnArgAspAla 519
QY 1109 AGCTCTCG-CCCTGCGAGCCCGCTGGGAGCATCCAGCAGCCCTGGAGCTCATCTTT 1167
DB |||||
DB 520 AsnLeuGlnAlaLeuIleAlaThrPheGlyAspIleAsnAlaAlaValGluArgLeuLeu 539
QY 1168 GCT 1170
DB |||||
DB 540 Ser 540

```

Search completed: November 12, 2003, 06:46:17
Job time : 55 secs

[illegible]

JS-09-735-251-3 (1-12'3) x US-09-890-688-86 (1-380)

QY 44 ATGTTCTCTCAGACTGCGCTGCGGTGAAGTGGCTGACACGACCTTACTTCAAG 103
DB 1 MetSerLeuSerAspTyrHisLeuAlaValLysLeuAlaAspGlnProLeuThrProLys 20
QY 104 TCTATTCTCGGTGCGCAGACAGCAATGGAGAAATACGGTAAAGGGGGTATAGTATT 163
DB 21 SerIleLeuArgLeuProGluThrGluLeuGlySerLeuGlySerLeuGlySerIle 40
QY 164 TCATTTCGAAGCAGCTTATTCCTGGTAACTCAGGAGTCTCTCCAGACCTGACCTG 223
DB 41 SerPheLeuLysGlnLeuIleAlaGlyLysLeuGlnGluSerValProAspProGluLeu 60
QY 224 ATTGATCTGATCTACTGTGGTGAAGCTAAAGATGACACGACCTTACTTATGGC 283
DB 61 IleAspLeuIleTyrCysGlyArgLysLeuLysAspAspGlnThrLeuAspPheTyrGly 80
QY 284 ATTCAAGCTGGTCACTGTCCATCTCTGCGAAGTCTGGCTGACCTGATCATCAAA 343
DB 81 IleGlnProGlySerThrValHisValLeuArgLysSerTyrProGluProAspGlnLys 120
QY 344 CGGAACCTGTGGCAAGTGGCTGATGAGAGTTCGGGTGTTCCACACTGGCCTG 403
DB 121 ProGluProValAspLysValAlaValMetArgGlnPheArgValLeuHisThrAlaLeu 120
QY 404 CACACGACTCTCTTACAGGAGAGGCTCTTACAGTCTGACCAATAGAGTCTCTG 463
DB 121 HisSerSerSerTyrArgGluAlaValPheLysMetLeuSerAsnLysGluSerLeu 140
QY 464 GATCAGATCATTTGGCCACCCCACTCACTCACTGACCTGATTCCTGGGTCTC 523
DB 141 AspGlnIleIleValAlaThrProLysSerSerAspProIleAlaLeuGlyValLeu 160
QY 524 CAGCAAGAGGACCT 583
DB 161 GluAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspThrLeuValPro 180
QY 584 GCTCAGCAGGCTGCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
DB 181 AlaHisProAlaLeuValAspAlaIleValLeuValLeuHisSerValAlaGlySerAla 200
QY 644 CCAATGCTGGGACTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
DB 201 ProMetProGlyThrAspSerSerSerSerSerSerSerSerSerSerSerSer 220
QY 704 CAGGTGCT 763
DB 221 ProGlyGlyPheLeuPheLeuGlyLeuMetArgSerLeuAlaAspPheHisProAsnThr 240
QY 764 AGGTCCACGCTCTACAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
DB 241 ArgSerThrProSerSerSerThrHisSerSerSerSerSerSerSerSerSerSer 260
QY 824 GCTGTGGGCCCCCGCCCATCAGCAAGTGAAGTGGCCACCGCTTGGCCCTGGCCAGC 883
DB 261 AlaAlaGlyProArgProIleThrGlnSerGlnAlaLeuAlaLeuAlaLeuAlaSer 280
QY 884 ACTCGGAGAGAGCTCTACACAAGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 943
DB 281 ThrProGluSerSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
QY 944 TCACCAATGCTCTCTGTTCTCAGTCCAGGAGCGCCATCAGCAATGATCTCTCTCAGCAA 1003
DB 301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAspAspLeuPheSerGln 320
QY 1004 GCCCTACAGCATGCCCTTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1063
DB 321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnTyrGlnPro 340
QY 1064 CAGCTGCGACGACTACGTGACATGCGGCATCCAGACAGTATGAGCTGCGGCCCTG 1122
DB 341 GlnLeuGlnGlnLeuArgAspMetGlyIleGlnAspAspGlnLeuSerLeuArgAlaLeu 360

QY 1123 CAGGCGACCGCTGGGACATCCAGACGACCTCGAGCTCATCTTTCTGGAGGACCCCA 1182
DB 361 GlnAlaThrGlyGlyAspIleGlnAlaAlaLeuGluLeuIlePheAlaGlyGlyAlaPro 380
RESULT 3
US-10-102-806-488
; Sequence 488, Application US/10102806
; Publication No. US2003054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: P033F1C1
; CURRENT APPLICATION NUMBER: US/10102,806
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2003-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 488
; LENGTH: 114
; TYPE: BPT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (95)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (113)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-488
Alignment Scores:
Pred. No.: 1,096-23 Length: 114
Score: 444.00 Matches: 92
Percent Similarity: 88.7% Conservative: 3
Best Local Similarity: 85.9% Mismatches: 10
Query Match: 19.2% Indels: 2
DB: 15 Gaps: 0
US-09-735-251-3 (1-125): x US-10-102-806-488 (1-114)
QY 8 GCCCAGAGAGAGAGAGAGCGCGAAAGAGAGAGAGATGCTCTCTCAGACTGGCACCTG 67
DB 10 GlyThrGlyArgGluArgGluArgGluArgMetSerLeuSerAspThrPheLeu 29
QY 68 GCGGTGAAGCTGGCTGACCCAGCCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 127
DB 30 AlaValLysLeuAlaAspGlnProLeuAlaProLysSerIleLeuGlnLeuProGluSer 49
QY 128 GAAGTGGAGATATCTCGCTAGCGGCTATAGTATTTCTTCTTCTGACACGCTTATGCT 187
DB 50 GluLeuGlyGlyTyrSerLeuGlyGlyTyrSerIleSerPheLeuLysGlnLeuIleAla 69
QY 188 GCGAAACTCCAGAGTCTGTTCTCCAGACCTGAGCTGATTGATCTGATCTGCTGCTGG 247
DB 70 GlyLysLeuGlnGluSerValProAspProGluLeuIleAspLeuIleTyrCysGlyArg 89
QY 248 AGCTTAAAGATGACACGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 307
DB 90 LysLeuLysAspAsp***ThrLeuThrSerThrValPheAsnLeuAlaPro-HisProC 109
QY 308 GTTCTCGGAAAGTCTCTG 324
DB 109 ysSer***Glu***Leu 114
RESULT 4

```

US-09-745-008-34
; Sequence 34, Application US/09745008
; Patent No. US20020137667A1
; GENERAL INFORMATION:
; APPLICANT: Chuenkova, Marina
; APPLICANT: Pereira, Miercio A.
; TITLE OF INVENTION: T. Cruzi Derived Neurotrophic Agents and
; FILE REFERENCE: Methods of Use Therefor
; CURRENT APPLICATION NUMBER: US/09745008
; PRIOR FILING DATE: 2000-12-23
; PRIOR APPLICATION NUMBER: US 60/172,481
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 1162
; TYPE: PRT
; ORGANISM: Trypanosoma cruzi
US-09-745-008-34

Alignment Scores:
Pred. No.: 0.00205 Length: 1162
Score: 163.50 Matches: 109
Percent Similarity: 36.16% Conservative: 43
Best Local Similarity: 24.94% Mismatches: 181
Query Match: 7.10% Indels: 98
DB: 18 Caps: 18

US-09-735-251-3 (1-1273) x US-09-745-008 34 (1-1162)
QY 59 TGGACCTGGCGGTGAAGCTGGTGGACACGACCTACTCCMAAGTCTATTCTCGGTG 118
DB 507 TyrosineValLeuThrMetAlaAsn
QY 119 CCAGACACACACTGGGAGAAATATCTGTAGGGGGCTATAGATTTCATTTCTGAGAG 178
DB 517 LysileGlySerValIleAspGly
QY 179 CTTATTCTCGCAACCTCCAGGATCTGTGTGACACCTGAGTGTATGAT 229
DB 229 LeuLeuGluGlySerGlyGlnThrValProAspGlyArgThrProAspLeuSerHis 537
QY 229 229
DB 538 PheTyValGlyGlyTyLysAlaSerAspValProThrIleMetThrValAsn 557
QY 230 CTGACTTACTTGGTTTGGTATAGTAAAGAAATAGACACTTACTTCTATGSC 293
DB 558 AsnValLeuLeuTyAspArgAlaIleAlaMetIleGluIleAlaGlyThrLeuPheLeu 577
QY 284 ATTCACACT-----GGTCCACTGTCCATGTTCTGGGAAAGTCTCGGCTGAACCTGAT 337
DB 578 SerGlnAspLeuIleGlyThrGlnAlaHisMetAspSerSerSerSerAlaHis 597
QY 338 CAGAAACCGGAA-----CTGTGGA AAGTGTGTTCCCATGAGAGAGTTCGGGTGTCAC 394
DB 598 SerThrProSerThrProAlaAspSerSerAlaHisSerThrProSerThrProValAsp 617
QY 395 ACTGCTCTGCACACACTCTCTCTATAGGAGGCGGTCTTAAAGATGCTCAGCAATAAG 454
DB 618 SerSerAlaHisSerThrProSer----- 625
QY 455 GAGTCTCTGGATCAGATCATTTGTGGCACCCCTCAGGCTCAGACGTGACCTATTGCTCT 514
DB 626 ThrProAlaAspSerSerAlaHisGlyThrProSerThrProValAspSerSerAlaHis 645
QY 515 GGGGTCTTCCAGGACAGGAGCTTTCTGTGTCTTCTGCTGATCCCAATATG-----CTGAT 571
DB 646 GlyThrProSerThrProAlaAspSerSerAlaHisGlyThrProSerThrProValAsp 665
QY 572 ACGTTGGTGTCTGCTCACCACGAGCTCTGTCATGCCATGCTGCTGTGCTGCTGCTGCT 611
DB 611

```

RESULT 5

```

; Sequence 20, Application US/09843746
; Publication No. US2003016650A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; APPLICANT: Honchell, Cynthia D.
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Mucin-Related Tumor Marker
; FILE REFERENCE: PG-0019 US
; CURRENT APPLICATION NUMBER: US/09/840,746
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Sub scrofa
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Genbank IT No. US2003016650A1 g915208
US-09-840-746-20

```

```

Alignment Scores:
Pred. No.: 0.00371 Length: 528
Score: 159.00 Matches: 91
Percent Similarity: 37.02% Conservative: 53

```

Best Local Similarity: 23.39% Mismatches: 132
Query Match: 6.91% Indels: 115
DB: 12 Gaps: 15

US-09-735-251-3 (1-1273) x US-09-840-246-20 (1-528)

```
QY 378 AGTCCGGGTTTGCACACTCCCTGACCA .....GCA 410
DB 24 SerSerSerValProProSerThrThrThrValGlnProSerSerSerGlySer 43
QY 411 GCTCTCTTACAGGAGGGGCTCTTACAGCTTCAGCACTACAGAGCTCTCGATCAGA 470
DB 44 AlaProThrThrSerAlaThrSerValGlnThrSerSerSerSer 59
QY 471 TCATTGTGGCCACCCAGCCCTCAATAGTACCTATTGCTCTGGGGTCTCCAGACA 530
DB 60 -----ProProSerSerThrThrThrSerValGlnThrSerSerSerSerVal 76
QY 531 AGGACCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 590
DB 77 ProThrThrSerThrThrSerValGlnPro-----Ser 87
QY 591 CAGCCCTCGTCAATGCCATCTGCTG-----GTCTGCACTCCGAGCAGGCGAT 640
DB 88 SerSerSerSerAlaProThrThrArgAlaThrSerValGlnSerSerSerSer 107
QY 641 GCGCCCAATGCTGGGACT -----CACTCTCTCTCCGAGCATGCTCTCCAGC 688
DB 108 AlaProSerSerThrThrSerValGlnProSerSerSerGlySerValProThrThr 127
QY 689 TCATACCGGATATGCCAGGTGGTTCTCTGTTTAAAGGCTCTAGATGATGAGGATGAC 748
DB 128 SerAlaThrSerValGlnSerSerSerSerSerAlaProThrThrSerAlaThrSer 147
QY 749 TTTCAACCAACACACAGGTCCAGACGC ----- 775
DB 148 ValGlnProSerSerSerSerSerProProSerSerThrThrValSerValGlnProSer 167
QY 776 ---TCTACAGTACTCCAGCTCCAGCCAGCTCCCTGGGATACAGTACAGTGTGGG 832
DB 168 SerSerSerSerAlaProThrThrSerAlaThrSerValGlnProSerSerSerSer 187
QY 833 CCC----- 835
DB 188 ProProSerSerThrThrValSerValGlnThrSerSerSerSerSerValProThrThr 207
QY 836 ---CGGCTCATCTCAAGAGTCAATGCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 890
DB 208 SerThrThrSerValGlnProSerSerSerSerSerValGlnThrSerAlaThrSer 227
QY 881 AGCACTCCGGAGAGAGCTCTCAAGATCAAGTCTCTGATC -----CAGGGTCTAT 931
DB 228 ValArgSerSerSerSerSerThrProProProSerThrThrSerValGlnProSer 247
QY 932 TCCTCAGGAGCTCACCAGTGTCTCTGGTCT-----CAGTCAGG 973
DB 248 SerSerSerSerAlaProThrThrSerAlaThrSerValGlnProSerSerSerSer 267
QY 974 ACGCCCATCAACCAATGA ----- 990
DB 268 ThrProSerProSerThrThrThrSerValGlnProSerSerSerSerAlaProThrThr 287
QY 991 -----TCTCTCAGCCAGAGCTCTACAGTCCCT-----TCAGGCTCTGGGCA 1035
DB 288 SerAlaThrSerValGlnProSerSerSerSerProProProSerThrThrThrSer 307
QY 1036 GCCCAGCTCTCAGAGCCAGTGCAGCCAGCTGCAGCAGTACGTGACATGGGCATCCA 1095
DB 308 ValGlnProSerSerSerSerSerProProThrThrThrThrThrSer-----ValGI 325
QY 1096 GGACATGAGCTGAGCTCGCGCCCTGTCAGG-----CCACCGGTGGGAGC 1140
DB 325 nProSerSerSerGlySerAlaProThrThrThrSerAlaThrSerValGlnProSerSerSe 345
```

RESULT 6

US-09-823-240-2

Sequence 2, Application US/C9823240

Patent No. US20020048913A1

GENERAL INFORMATION:

APPLICANT: Frank B. Gertler

APPLICANT: James E. Bear

APPLICANT: Jürgen Wehland

APPLICANT: Joseph Loureiro

TITLE OF INVENTION: Methods and Products for Regulating Cell

FILE OF INVENTION: Mobility

FILE REFERENCE: WO656/7064 (HCL)

CURRENT APPLICATION NUMBER: US/09/823,240

CURRENT FILING DATE: 2001-03-30

PRIOR APPLICATION NUMBER: 60/194,564

PRIOR FILING DATE: 2000-04-03

NUMBER OF SEQ ID NOS: 11

SOFTWARE: FastSeq for Windows Version: 3.0

SEQ ID NO 2

LENGTH: 902

TYPE: PRT

ORGANISM: Mus musculus

US-09-823-240-2

Alignment Scores:

Pred. No.: 0.00401 Length: 802
Score: 159.00 Matches: 117
Percent Similarity: 35.36% Conservative: 44
Best Local Similarity: 25.71% Mismatches: 171
Query Match: 6.91% Indels: 123
DB: 9 Gaps: 23

US-09-735-251-3 (1-1273) x US-09-823-240-2 (1-802)

```
QY 2 CGAGAGGGTCCAGAGAGAGAGAGAGAGCGC-----GAAAGAGAGAGATGCT 49
DB 228 ArgGlnArgGlnGluArgGluArgGluArgGluArgGluArgGluArgGluVal--- 246
QY 50 CTCTCAGACTGCCACTCGCGGTGAAGCTGAGCTGAGCTGACAGCCACTTACTCCAAAGTCTATT 109
DB 247 -----GluTrpGluArgGluArgMetSerAsn-----AlaAlaProSerSerAsp 262
QY 110 CTTCGGTTCAGAGAGAGAGAGAGAGAGTGGAGAAATCTCGCTA-----GGGGGCTAT 157
DB 263 SerSerSerSerAlaProLeuProLeuProLeuProLeuProLeuProLeuProLeuPro 282
QY 158 AGTATTTCTATTTCTCAAGCAGCTATTCTCGCAAACTCCAGGAGCTGTGTTCCAGAGCCCT 217
DB 283 ProProSerSerAlaLys-----ValIleSerAlaProValSerAspAlaThrProAspTyr 301
QY 218 GAGCTGATGATCTGATCTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 277
DB 302 AlaValValThrAlaLeu----- 307
QY 278 TATGCAATTCACCTGGGTCCACTCTCTCCATGTTCTGCGAAAGTCTGCTGCTGAACCTGAT 337
DB 308 -----ProProThrSer 311
QY 338 CAGAAACCGAACTCTGTGCAAAAGTGGCTGCCATGACAGAGTTCGGGTGTTGCACACT 397
DB ----- 397
```

```

Db 312 ThrProThrProProLeuArgHisAlaAlaThrArgPheAlaThrSerLeuGlySer 331
QY 398 GCGCTCCAGCAGCAGCTCTCTTACAGGAGGGGCTCTTAAGATCTCAGCAATAAGAG 457
Db 332 AlaPheHisProValLeuProHisThrAlaThrValProArgProLeuAshLysAsn 350
QY 458 TCTCTGGATCAGATCATTTGGCCACCCGAGGCTTCAGCAGTACCTCTCTTGGG 517
Db 351 SerArgProSerProValAsnThrProSerSerGlnProAlaAlaLys 368
QY 518 GTTCTCCAGGACAAGGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 577
Db 368 sSerCysAlaThrProThrSerArgPheSerProLeu 380
QY 578 GTGCTGTCTCACCCAGCCCTCTTAATGCCATTGTCTCTCTCTCTCTCTCTCTCT 637
Db 381 ProProSerProThrAlaProProSerProThrAlaProProSerProThrAla 395
QY 638 A-----GTGCTCAATGCTCTGGAGTCTGACTCTCTCTCTCTCTCTCTCTCT 679
Db 395 aThrGlyProArgProValLeuPheValCysValSerSerProValProGlnMe 413
QY 680 CCTCCAGCTCATCGGGGATATTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 739
Db 413 tProProSerProThrAlaProProSerProThrAlaProProSerProThrAla 425
QY 740 GAGGATGACTTTCACCCCAACATAGGTCTATCTCTCTCTCTCTCTCTCTCTCTCT 799
Db 425 pSerValThrThrProValSerProProThrSerGlyProAlaAlaProProThrPr 445
QY 800 CAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 859
Db 445 oProProProProProProProProProProProProProProProProProPro 457
QY 860 GCGAGCGCTTGGCTGCGGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 904
Db 457 uProProProProProProProProProProProProProProProProProPro 475
QY 905 ACAGCGACTCTCTGCGCAGGAGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 955
Db 476 SerProProGlyThrProLeAlaSerProSerSerLeuProSerValLeuPro 495
QY 956 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 997
Db 496 SerProSerAlaGlyValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 515
QY 998 AATTAAGCCCTACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1057
Db 516 SerSerAlaSerGlyProGlyLeAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 533
QY 1058 CAGCCCGAGCTGCGCAGCTACGATGATGATGATGATGATGATGATGATGATGATG 1117
Db 534 ThrProGlnProProProProProProProProProProProProProProPro 541
QY 1118 CCGTGCAGG-----CCACCGATGGGACATTCAGAGAGGCTCTGAGGCTCTCT 1171
Db 542 ProProAlaProProProProProProProProProProProProProProPro 561
QY 1172 GAGGAGCCCATGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1231
Db 562 ProProProProGlyProProProProProProProProProProProProPro 579
QY 1232 CTTGGGAGGCACTCATGAAGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1290
Db 580 ProProProProProProProProProProProProProProProProProPro 599

```

RESULT 7

US-10-213-509-5

Sequence 5, Application US/10213509

Publication No. US20030054485A1

GENERAL INFORMATION:

APPLICANT: Weiss, Joseph

APPLICANT: Scott, Matthew

TITLE OF INVENTION: JELLY BELLY GENES AND THEIR USES

FILE REFERENCE: STAN-232

CURRENT APPLICATION NUMBER: US/10/213,509

CURRENT FILING DATE: 2002-08-06

PRIOR APPLICATION NUMBER: 63/311,720

PRIOR FILING DATE: 2001-08-09

NUMBER OF SEQ ID NOS: 5

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 5

LENGTH: 4123

TYPE: PRT

ORGANISM: H. sapiens

US-10-213-509-5

Alignment Scores:

Pred. No.:	0,0114	Length:	4123
Score:	154.50	Matches:	116
Percent Similarity:	28.03%	Conservative:	25
Best Local Similarity:	23.06%	Mismatches:	144
Query Match:	6.71%	Indels:	219
DB:	15	Gaps:	30

US-09-735-251-3 (1-1293) x US-10-213-509-5 (1-4123)

```

QY 196 CCAGAGTCTGTTCCAGACCCCTGAGCTGATGTGATCTACTGTGTCGGAAGTAA 255
Db 2773 ProGlyGlyCysProLeuProCysGly-----TrpSer----- 2783
QY 256 AGATGACCACACACTTGACTTCTATGGCATTCAACCTGGGTCCACTGCTCCATG 315
Db 2784 -----AlaTrp-----SerSerTrpAlaProCysAspArgSer--- 2794
QY 316 AAAGTCTGCTGCTGAACCTGATCAGAAACCGAACCTGTGGACAAAGTGGCTGCC 375
Db 2795 -----CysGly-----SerGlyValArgAla 2801
QY 376 AGAGTTCG-----GGTTTCCACACTGCCCTCCGACAG 408
Db 2802 ArgPheArgSerProSerAsnProAlaAlaTrpGlyAlaProCysGluGlyAsp 2821
QY 409 CAGTCTCTCTTACAGGA-----GGGGTCTTTAAGATGCTCAGCAATAA 453
Db 2822 ArgGlnGluLeuGlnGlyCysHisThrValCysGlyThr----- 2834
QY 454 GGAGTCTGTCGATCAGATCATCTGGGACCCAGGCTCAGCAGTGACCTATTGG--- 510
Db 2835 GlyIleAlaGlySerLeuGlyAlaGlyValPro-----ProSerSerGlnPheCysThr 2853
QY 511 -----TCTTGGGGT----- 519
Db 2854 LeuArgThrHisGlyMetGlyProThrAspHisSerThrTrpGlyIleGluValPheGly 2873
QY 520 -----TCTCCAGGACAAGGA 534
Db 2874 TrpThrProTrpThrSerTrpSerSerCysSerGlnSerCysLeuAlaProGlyGly 2893
QY 535 CCT-----CTTCTCTGCTTCTGCTGATCCCAATATGCTTGTATACGTT 576
Db 2894 ProGlyTrpArgSerArgSerArgLeuCys----- 2903
QY 577 GGTGCTGCTCACCAGCCCTCTGTCATGTCATGTCCTGTTCTGCTCCTCCTAGCAG 636
Db 2904 -----ProSerProGlyAspSerSerCysProGlyAspAlaThrGlnGlu 2919
QY 637 CAGTGCCC-----CAATGCTGGGACTCCTCTCTCCCGAGGAGCATGCCCTCCA 686
Db 2920 ProCysSerProProIleGluCysThrGlyPheCysAlaProGlyCysThrCysProPro 2939
QY 686 ----- 686
Db 2940 GlyLeuHisLeuHisAsnAlaSerCysLeuProArgSerGlnCysProCysGlnLeuHis 2959
QY 687 -----GCTCATACCGGGATATGCCAGGTTGCTCTCTCT 719

```

```
Db 2960 GlyGlnLeuTyrAlaSerGlyAlaMetAlaArgLeuAspSerCysAsnAspCysThrCys 2979
QY 720 TTGAAGGCTCTCAGATG-----
Db 2980 ValSerGlyLysMetAlaCysThrSerGluArgCysProValAlaCysGlyTyrSerPro 2999
QY 738 -----
Db 3000 TrpThrLeuTrpSerLeuCysSerLysSerCysAsnValGlyIleArgAlaArgPheAla 3019
QY 756 CAACACACAGGTCACACCCCTCTCAAGATATCCAGATCCCGCCAGGCTCCCTGGGGT 815
Db 3020 AlaGlyThrAlaPro---ProAlaAlaPheGlyGlyAlaGlyCysGlyGlyProThrMet 3038
QY 816 ACATGGAGCTGCTGGG-----
Db 3039 GluAlaGluPheCysSerLeuArgProCysProGlyProValProGlyMetCys---Pro 3057
QY 864 CGGCTTGGCCCTGG---CCAGCAATCCGAGAGCAGCTCTCACACACCGACTCTGGCA 920
Db 3058 ArgAspLysGlnTrpLeuAspCysAlaGlnGlyProAlaSerCysAlaGluLeuSerAla 3077
QY 921 CCAGGGTCATTCTCCAGGAGCCCTGAC-----CAATGCTCT----- 956
Db 3078 ProArgGlyThrAsnGlnThrCysHisProGlyCysHisCysProSerGlyMetLeuLeu 3097
QY 957 CTGGTGTCCAGTCAGGAGCGCAACCAATATCTCTTTCAGCCAGGCTTACAGCATG 1016
Db 3098 LeuValSerProArgGlyHisProGlyProLeuGlyAlaSerValGlnPro----- 3114
QY 1017 CCCTTCAGGCTCTCCGGCAGCGCAACCTTCAGAGCAGTGGACAGT---CCAGGCTGACG 1073
Db 3115 ---ProValAlaLeuProGlyAlaIleGlyThrGlySerValProGlyAlaGly 3131
QY 1074 AGCTACGTGACATGGCATCCAGACGATGATGAGCTCCGCTCCGCTCCGCTCCGCTCCG 1133
Db 3132 GlyTrpGlyProTrp----- 3138
QY 1134 TGGGAGATCCAGC---AGGCTCGACTCATCTT-----TGCCTG 1172
Db 3139 TrpSerHisCysSerArgSerCysGlyGlyLeuArgSerAlaThrArgAlaCysAsp 3158
QY 1173 AGAGCCCATCAACTCCCTG-----CTGCTGCTGAACCCCGCCAGCAAGTT 1217
Db 3158 GlnProProGlnGlyLeuGlySerProGlyGlyProGlyProGlyProGlyProGly 3178
QY 1218 GCAGAGGCTACTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 1255
Db 3179 CysGlnAlaLeuProCysProValAlaMetAlaArgLeuLeuAlaGlyAlaGlyAla 3198
QY 1266 TCCCTGT 1272
Db 3198 TrpCys 3200
```

RESULT 8

```
US-10-146-473-47
Sequence 47, Application US/10146473
Publication No. US2003010888A1
GENERAL INFORMATION:
APPLICANT: Scanlan, Matthew
APPLICANT: Gout, Ivan
APPLICANT: Stockert, Elisabeth
APPLICANT: Gure, Ali
APPLICANT: Chen, Yao-Tseng
APPLICANT: Old, Lloyd
TITLE OF INVENTION: Breast Cancer Antigens
FILE REFERENCE: L00461/70130(JRV)
CURRENT APPLICATION NUMBER: US/10146473
CURRENT FILING DATE: 2002-05-15
PRIOR APPLICATION NUMBER: US 60/291 156
PRIOR FILING DATE: 2001-05-15
NUMBER OF SEQ ID NOS: 82
```

```
SOFTWARE: Patentin version 3.0
SEQ ID NO 47
LENGTH: 624
TYPE: PRT
ORGANISM: Homo sapiens
US-10-146-473-47
```

```
Alignment Scores:
Pred. No.: 0 0112 Length: 624
Score: 152.50 Matches: 97
Percent Similarity: 14.04% Conservative: 70
Best Local Similarity: 13.76% Mismatches: 177
Query Match: 6.62% Indels: 147
DB: 15 Gaps: 19
```

```
US-09-735-251-3 (1-1213) x US-10-146-473-47 (1-624)
```

```
QY 80 GTGACACGACCTTACTCCAAAGTCTATT-----CTTCGTTCCAGACACAGAA 130
Db 24 AlaAlaAlaProAlaGlyProLysIleIleLysValThrValLysThrProLys---Glu 42
QY 131 CTGGGAGATATCTCGCTA---GGGGCTATAGTATTTCATTTCAGACAGCTTATTCT 187
Db 43 LysGluGluPheAlaValProGluAsnSerSerValGlnGlnPheLysGluAlaIleSer 62
QY 188 GGCACAACTCCAGGAGTCTGTTCACACCCCTGAGCTGATTGATCTCTACTGTGGTCGG 247
Db 63 LysArgPheLysSerGlnThr-----AspGlnLeuValLeuIlePheAlaGlyLys 79
QY 248 AGCTAAAGATGACACAGACACTTGCCTTCTATGACATTCACCTGGGTCCACTTCAT 307
Db 80 IleLeuLysAspGlnAspThrLeuIleGlnHisGlyIleHisAspGlyLeuThrValHis 99
QY 308 GTCTGCGAACTCTGGCTCAACTGAT---CAGAAACCGGAACCTGTGCACAAAGTG 364
Db 100 LeuValLysSerGlnAsnArgProGlnGlnSerThrGlnProSerAsnAlaAla 119
QY 365 GCTGCCATAGAGAGTTCCGGGTGTGCACACTGCCCTGCACAGCAGCTCTCTTACAGG 424
Db 120 Gly-----ThrAsnThrSerAlaSerThrProArg 130
QY 425 GAGCGCTTTTAAGATGCTCAGCAAT----- 451
Db 131 SerAsnSerThrProIleSerThrAsnSerAsnProPheGlyLeuGlySerLeuGlyGly 150
QY 452 -----AAGCAGTCT 460
Db 151 LeuAlaGlySerSerSerLeuGlyLeuSerSerThrAsnPheSerGluLeuGlnSerGln 170
QY 461 CTGGATAGATATATGATGCCACCCCA-----GGCCTCAGCAGTACCCCTATT 508
Db 171 MetGlnGlnGlnLeuMetAlaSerProGluMetMetIleGlnIleMetGluAsnProPhe 190
QY 509 GCTCTTGGGGTCTTCACAGACAGGACCTCTCTCT-----GTCTTCCTGATCCCAAT 562
Db 191 ValGlnSerSerLeuSerAsnProAspLeuMetArgGlnLeuIleMetAlaAsnProGln 210
QY 563 ATGCTTGAACGTTGTGGCTGCTCTCACCCAGCCCTCTGCAATGCC----- 607
Db 211 Met---GlnGlnLeuIleGlnArgAsnProGluIleSerHisLeuLeuAsnAsnProAsp 229
QY 608 ATTGTCTCTGTTCTGCACCTCCAGTACGAGCAGTCCCGCAATGCCCTGGGACTGCTCTCT 667
Db 230 IleMetArgGlnThrLeuGluIleAlaArgAsnProAlaMetMetGlnGluMetMetArg 249
QY 668 TCCCGGAGCATGCCCTCCAGCTCATACCCGGGATATGCCAGGTGGGTCTC---CTGTTTCAA 724
Db 250 AsnGlnAspLeuAlaLeuSerAsnLeuGluSerIleProGlyGlyTyrAsnAlaLeuArg 269
QY 725 GGGCTCTCAGATGATGAGTACCTTACCCAAACACCCAGGTCCACACCCCTCTAGCAGT 784
Db 270 ArgMetTyrThrAspIleGlnGlnProMetLeuAsnAlaAlaGlnGlnPheGlyGly 289
```



```
Db 404 rAlaProLeuProGlnProAlaSerAlaProGlySerAlaProGlyProAlaGlnXc 424
Qy 843 -----TCACCCAGAGTGGCTGG----- 860
Db 424 tProLeuProAspAlaThrProSerArgSerPheThrMetSerAspThrGlySerPh 444
Qy 861 -----CCACCGCTTGGCCCTGCCCAGCACTGGGAGACCACTCTCCACACCCAGCT 913
Db 444 eHisLeuProProArgProSerThrProValArg----- 456
Qy 914 CTGGCACCAGGGTCACTCTCCACAGGACCTCAATGCTCTGGTCCAGTCAGGG 973
Db 457 -----ProGlnProLeuProGlnArgPro----- 464
Qy 974 AGCCCATCACCAGTATCTCTTACGCAAGCC-----TACAGCATGCTCTCAG 1024
Db 465 -----ProThrProAlaSerValAlaLeuProSerProAlaThrAlaProHisLeuGl 483
Qy 1025 GCTCTGGGAGCCAGCCCTCAAGCCAGTGGAGCCAGCCAGCCAGCTGACAGCTACGTGAC 1084
Db 483 uProThrArgValProSerProAlaGProSerAlaAlaProAlaAlaSerValGlnHisH 503
Qy 1085 ATGGCATCCAGGAGCAAGCTGAGCTGAGCTGCGCCCTGCAGCCAGCCGGGGGACATCC 1144
Db 503 lAlaThrAlaSerAlaGlyAlaHisThrAlaProAlaAlaGlnValHisThrArg 521
Qy 1145 AAGCAGCCCTGGAGCTGATCTTTCTGGAGGAGCCCACTGAATCCCTCTCCCTGAA 1204
Db 522 ProGln-----AlaThrArgArgArgArgAlaGArgProLysArgProGly 537
Qy 1205 CCCCAGCA 1213
Db 538 ProProAla 540

RESULT 10
US-09-893-519A-14
; Sequence 14, Application US/0989351A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMESON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BUCMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Tharara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOVARNITSKY, Svetlana
; APPLICANT: MENDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: MCCOY, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HAQ, Tariq
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAVIDOV, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/10548-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1023
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Corresponds to SEQ ID NO: 87
```

```
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Human Genbank/CAA72189
; DATABASE ENTRY DATE: 1997-06-25
; RELEVANT RESIDUES: (1)..(1023)
US-09-893-519A-14

Alignment Scores:
Pred. No.: C.0257 Length: 1023
Score: 148.00 Matches: 91
Percent Similarity: 31.21% Conservative: 17
Best Local Similarity: 26.30% Mismatches: 134
Query Match: 6.43% Indels: 104
DB: 11 Gaps: 15

US-09-735-251-3 (1-1273) x US-09-893-519A-14 (1-1023)
Qy 355 GCACAAAGTGGTCCATGAGAGAGTTCGGGTGTTGCACACTGCCCTGCACAGCAGCTC 414
Db 74 GlyProAlaAlaProAlaGluGlyAlaProGlyAlaAlaProGluProProAlaGly 93
Qy 415 CTCTTACAGGGAGGGCTCTTTAAGATGCTCAGCAATAAGGAGTCTCTGGATCAGATCAT 474
Db 94 ArgAlaArgProGlyGly-----GlyGlyProGlnArg 104
Qy 475 TGTGGCCACCCAGCCCTCAGCAG-----TGACCTATTGCC 510
Db 105 ProGlyProProSerProArgArgProLeuValProAlaGlyProAlaProProAlaAla 124
Qy 511 TCTTGGGTCTCCAGGACCAAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTGA 570
Db 125 LysLeuArgProProProGluGlySer----- 133
Qy 571 TAGTTGGTGGCTGTCTACCCAGCCCTGCTCAATGCCATTGTC----- 613
Db 134 ---AlaGlyAlaCysAlaProValProAlaAlaAlaValAlaAlaGlyProGluPr 152
Qy 614 -----CTGGTTCTGCACCTCGTAGCAGGCGAG 639
Db 152 oAlaProAlaGlyProAlaLysProAlaGlyProAlaGlyProAlaAlaAlaArgAlaGlyPr 172
Qy 640 TGCCCAATGCTGGGAGTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTATACCGGGA 699
Db 72 OglyProGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyProGlyA 192
Qy 700 TATGCCAGGTGGCTTCTCTTTTGAAGGGCTCT---CAGATGATGAGGATGACTTTTCAACC 756
Db 192 LaAlaGlnThr-----LeuAsnGlySerAlaAlaLeuLeuAsnSerHisAlaAla 209
Qy 757 AAACATCA-----GGTCCACACCCCTCTAGCAGTACTCTCCAGCTC 795
Db 209 LaAlaProAlaValSerLeuValAsnGlyProAlaAlaLeuLeuProLeuProLysP 229
Qy 796 CCGCCAGCTCTCTGGGTACAGTGGAGCTCTCTGGGCC-----G 837
Db 229 roAlaAlaPro-----GlyThrValIleGlnThrProPheValGlyAlaAlaAlap 247
Qy 838 GCCATCACCCAGAGTGAAGCTGGCCACCCCTTGGCCCTGGCCAGCACTCCCGAGAGCGAG 897
Db 247 roProAlaProAlaAlaProSerProAlaAlaProAlaProAla----- 262
Qy 898 CTCTCACACACCGACTCTCTGGCAGCCAGGGTCACTCTCTCAGGAGCACTCCACCAATGTCCTC 957
Db 263 -----AlaProAlaAlaAlaProProProProProProAlaProPro- 275
Qy 958 TGGTGTCCAGTCAGGACCCCAATCATCACCATGATCTCTTTCAGCCCAAGCCCTACAGCATGC 1017
Db 276 -----AlaThrLeuAlaArgPro-----P 282
Qy 1018 CTTTACAGGCTCTGGCCAGCCAGCCAGCTTTCAGAGCCAGTGGAGCCAGCTCAGCAGCT 1077
Db 282 roGlyHisProAlaGlyProProThrAlaAlaProAlaAlaValProProProAlaAla 301
Qy 1078 ACCTGACATGGGCATCCAGGACGATGAGCTGAGCTGAGCTGGCGCTCTCGCGCTCTCGAGG---CCACCGGT 1134
```

```
Db 302 GlnAsnGlyGlySerAlaGlyAlaAlaProAlaProAlaProAlaAlaGlyGlyProAla 321
QY 1135 GGGGACATCCAGAGCCCTGGATCTCATCTCTCTGGAGGAGCCCATCACTCCCTGG 1194
Db 321 aglyValSerGlyGlnPro 334
QY 1195 TTCCCTCAACCCAGCCAGGAGTTGAGAGCTACTGCGCTTGGAGGAGCATGAAGGT 1254
Db 334 alaProAlaPro 355
QY 1255 GCTCCATCT 1264
Db 352 aProProAla 355

RESULT 11
US-10-156-761-11093
; Sequence 11093, Application US/015-761
; Publication No. US2003019015A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/015-761
; PRIOR FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-04089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-72697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11093
; LENGTH: 903
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11093

Alignment Scores:
Pred. No.: 0.0296 Length: 903
Score: 147.00 Matches: 81
Percent Similarity: 37.38% Conservative: 33
Best Local Similarity: 26.56%
Query Match: 6.39%
DB: 15

US-09-735-251-3 (1-1273) x US-10-156-761-11093 (1-903)
QY 551 GGTATCCCAATATGTTTATATG 903
Db 91 AlaAspAspValGlnAlaAlaProGlnAspSerValProGlyProPro 108
QY 599 GTCAATGCCATTGCTGCTGCTCC KATCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 658
Db 109 109
QY 659 GACTCCTCTTCC 703
Db 118 AspThrAsnAlaGluAlaGlyGlnAspArgSerValProLeuProAlaThrValPheAla 137
QY 704 CCAGGTGGCTTCTGTTTGAAGAGCTCTCAGATGATGAGATGACTTTACCAACACACC 763
Db 138 ProProPro 152
QY 764 AGGTCCACACCCCTCTAGCAGTACTCCAGCTCCGCGCCAGCTCCCTGGGGTACAGTGA 823
Db 153 ProProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 167
QY 824 GCTCTGGGCCCCGGGCCCATCACTCAGAGTGAAGTGGCCACCGCCTTG 877
```

```
168 GlySerGly 184
QY 878 GCCAGCACTCGGAGAGCAGCTCTCACACAGCCACTCTCTGGC----- 919
Db 185 ProAlaAlaProGlnGlyProVal-----ThrProProGlyProGlyAlaProArgVal 203
QY 919 ----- 919
Db 204 ProSerGlyProSerAlaAlaGlnGlyProGlyAlaProGluGlyAlaAlaProSerAla 223
QY 920 ----ACACAGGCTCATCTCTCAGGAGCTCTACCAATGCTCTCTGTTCCAGTCAGGAGC 976
Db 224 ProSerGlyGlyProGlnGlyProVal-----ThrProProGlyProGlyAlaProSerGly 243
QY 977 CCA-----TCACCATGATCTCTTC 997
Db 243 ProGlnGlyProGlyAlaValProGlyThrProProProGlyAlaProGlnAlaProGl 263
QY 998 AGCAAGCCTTACAGC-----ATGCCCTTCCAGCTCTCGGCGAGCCCGCCCTTCAGAGC 1051
Db 263 YAla- ProHisAlaProGlyAlaProGlnAlaProThrGlySerPro-----Va 279
QY 1052 CAGTGGCAGCCCGCAGCTGCGAGCAGCTACGTGACATGGGCATCCAGGAGCATGAGCTGAGC 1111
Db 279 lProGlyArgProLeuAlaProAsnAlaGlyAspIleAlaAspAlaAlaThrSerLyAl 299
QY 1112 CTGGGCGCTCGAGCGCCAGCGTGGGGACATCCAAAGCAGCGCTCGGAGCTCATCTCTTCTG 1171
Db 299 aAlaProArgGlyAlaAlaAlaGlyAlaProArgProGlyAlaProGlyValP 319
QY 1172 GAGGAGCCCATCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1216
Db 319 roGlyAlaProGlyThr-ProGlyAlaProGlyAlaMetProProAlaSerGlyProGly 338
QY 1217 -----TCAGAGGCTACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1261
Db 339 AlaProGlyThrProAlaGlyGlyThrValProThrGlnLeuValSerGlnLeuGlyPro 358
QY 1262 TCTCTCTCT 1270
Db 359 AspGlyPro 361

RESULT 12
US-10-146-473-50
; Sequence 50, Application US/0146473
; Publication No. US2003010888A1
; GENERAL INFORMATION:
; APPLICANT: Scanlan, Matthew
; APPLICANT: Gout, Ivan
; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tseung
; APPLICANT: Old, Lucy
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: US461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146.473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 50
; LENGTH: 2971
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-473-50

Alignment Scores:
Pred. No.: 0.0474 Length: 2971
Score: 145.50 Matches: 118
Percent Similarity: 33.80% Conservative: 50
Best Local Similarity: 23.74% Mismatches: 181
```


[illegible]

Db	406	SerLeuSerGlnAsnProAspLeuAlaAlaGlnMetMetLeuAsnAsnProLeuPheAla	425
Qy	854	-----GAGCTGGCCACCGCTTGGCCCTGGCC	880
Db	426	GlyAsnProGlnLeuGlnGlnMetArgGlnGlnLeuProThrPheLeuGlnGlnMet	445
Qy	881	AGCACTCGGAGAGCAGCTCTCACACCGACTCCT	916
Db	446	GlnAsnProAspThrLeuSerAlaMetSerAsnProArgAlaMetGlnAlaLeuLeuGln	465
Qy	916	-----	916
Db	466	IleGlnGlnGlyLeuGlnThrLeuAlaThrGlnAlaProGlyLeuIleProGlyPheThr	485
Qy	917	-----GGCACCCAGGTTCATCTCCAGGGACCTCACCAATGTCCTCT	958
Db	486	ProGlyLeuGlyAlaLeuGlySerThrGlyGlySerSerGlyThrAsnGlySerAsnAla	505
Qy	959	GGTGTCCAGTCAGGGACGGCC-----ATCCACCAATGATCTCTTCAGCCACGAC	1006
Db	506	ThrProSerGlnAsnThrSerProThrAlaGlyThrThrGlnProGlyHisGlnGlnPhe	525
Qy	1007	CTACAGATGCCCTTCAGGCCTCTGGG-----CAGCCACCGCTTCAGAGC	1051
Db	526	IleGlnGlnMetLeuGlnAlaLeuAlaGlyValAsnProGlnLeuGlnAsnProGluVal	545
Qy	1052	CAGTGCAGCCCCAGCTGCAGCAGCTACGTGACATGGGCATCCAGGACCATGAGCTGAGC	1111
Db	546	ArgPheGlnGlnLeuGlnLeuSerAlaMetGlyPheLeuAsnArgGluAlaAsn	565
Qy	1112	CTCGC-GCCCTGCAGGCCACCGGTGGGGCATCCAAAGCAGCCCTGGAGCTCATCTTGGCT	1170
Db	566	LeuGlnAlaLeuIleAlaThrGlyGlyAspIleAsnAlaAlaIleGluArgLeuLeu	584
Qy	1171	GGAGGAGGCCCA	1182
Db	585	GlySerGlnPro	598

49 GCGG-GAAGCTGGCTGACCAAGCCACCTACTCCAAAGTCTATTCTTCGGTGGCCAGAGACA 127

Best Local Similarity: 25.67% Mismatches: 86
Query Match: 6.21% Indels: 136
DB: 12 Gaps: 17

US-09-735-251-3 (1-1273) x US-10-017-151-1350 (1-727)

```
QY 326 CQTGAACCTGATCAGAAACGGGAATTTTGGAAAGTGGCTGGCATGAGAGA-----G 379
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 133 ProGlnProArgGluProSerPro-TripSerProSerValPro-GlyAsnProSerV 152
QY 380 TTCCGGGTGTTCCACACTGCCCTACAGACAGTCTCTTACAGAGAGAGGCTTTTAAAG 439
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 152 alProG.YThrProSerIlePro----- 159
QY 440 ATGCTCAGCAATAGAGAGTCTCTACATCAGATATTGGGGACGCC-----A 487
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 160 -----GlyThrProSerSerGlyAsnProSerValProGlyT 172
QY 488 GGCCTCAGCAGTGACCTATTCTTTTGGGT----- 519
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 172 hrProGln-----SerTrpLysProLeuSerProGlyAsnProSerValA 187
QY 520 -----TCTCAGAGCAAGAGCTCTC----- 541
Db 187 rgLysProLeuSerProG.Y.LysProLeuSerProArgLysProLeuSerProGlyAsn 206
QY 542 -- TCTGCTTCGGTCATACCAATATGCTTGATAGTTCGGTGGCTGCTACCCAGGCCCTC 598
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 207 ProSerValProG.YThrProGlnSerGlyAsnProSerValProGlyAsnPro----- 224
QY 599 GTCAATGCCATTGTCTCTCTTCTTATTCGGTAGAGACATGACCCCAATCTCTGGGACT 658
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 225 -----SerValProGlyAsnProSerValProGlyThr 235
QY 659 GACTCTCTTCCGGAGGATGATCCAGCTCATACGGGATATGCCAGTGGCTTCTCTG 718
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 236 ---ProGlnSerArgGluProGlnSerGlyAsnProSerValProGly----- 251
QY 719 TTTGAAGGGCTCTCAGATGATGATATGACTTTACTTAAACATCAGGACATACCCCTCT 778
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 252 -----ThrProGlnSerGlnIleThrProGln 260
QY 779 ASCAGTACTCCAGCTCTGGCTAGAGCTGCTGCTGGGTACAGTGGAGCTGCTGGCCCGG 838
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 261 SerArgLysProLeuSerProArgLysnProSerProArgLysnProLeuGlyProG 280
QY 839 C-----CATTAAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 883
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 280 yAsnProLeuSerGlnIleThrProGlnSerGlyAsnProSerValProGlyAsnPro 300
QY 884 ACTCGGAGAGCAGCTCTACAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 943
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 300 rValProArgAspProLeuSer-----ProGlyAsnPro 311
QY 944 TCACCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 311 o-----SerAlaProG.YThrProGlnSerLeuGluProLeuSerPro 325
QY 998 AGCCAGGCCCTACAGC-----ATGCCCTCAAGCTCTGGCCAGCCAGCTTCAGAGC 1051
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 325 ogLysProLeuSerProGlyAsnProPheAsnPro-----GlyAsnProSerValArgS 344
QY 1052 CAGTGGCAGCCCGCAGCTGAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1111
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 344 sProLeuSerPro----- 348
QY 1112 CTGGGGCCCTGCGAGGCCACCGCTGGGACATCAGAGCCCTGGAGCTCATCTTTGCTG 1171
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 349 -----GlyAsnProLeuSerProGlyLysProLeu----- 358
QY 1172 GAGGAGCCCATGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1208
   |||||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 359 ----SerProArgAsnProSerValProGlyAsnPro 369
```

Search completed: November 12, 2003, 07:00:40
JCB time : 64.5 secs

SEQUENCE CHARACTERISTICS:
 LENGTH: 410 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-725-758A-4

Alignment Scores:

Pred. No.: 1,42e 182 Length: 410
 Score: 2133.00 Matches: 410
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 92.66% Indels: 0
 Gaps: 0

US-09-735-251-3 (1-1273) x US-08-725-758A 4 (1 410):

```

QY 44 ATGCTCTCTCAGACGGACCTTGGCGTGAAGCTGGCTGACGACGCACTTACTCTCAAG 103
DB 1 MetSerLeuSerAspTrpHisLeuAlaValLysLeuAlaAspGlnProLeuTrpOlys 20
QY 104 TCTATTCTTCGGTTCCAGACAGCACTGGGAGAAATACCTGGTGGGAGCTATATATT 163
DB 21 SerLeuArgLeuProGluTrpGluLeuGlyGlnTyrSerLeuGlyGlyTyrSerLeu 40
QY 164 TCATTCTCAAGCAGCTTATCTGGGCAACACCCAGGAGCTGTTTCCCAACCTGAGCG 223
DB 41 SerPheLeuLysGlnLeuLeuAlaGlyLysLeuGlnLysValProSerProGluLeu 60
QY 224 ATTGATCTGATCTACTGTGTGTAAGCTAAAGATGACGACGACCTTACTCTCATGCG 283
DB 61 IleAspLeuIleTyrCysGlyAlaLysLeuLysAspAspGlnTrpLeuAspPheTyrGly 80
QY 284 ATTCACCTGGTCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 343
DB 81 IleGlnProGlySerThrValHisValLeuArgLysSerTrpProGluProAspGlnLys 100
QY 344 CCGGACCTGTGCACAAAGTGTTGTCATGACGAGAGCTCCGGTGTGTGACACTGCCGT 403
DB 101 ProGluProValAspLysValAlaAlaValArgGluPheArgValLeuHisThrAlaLeu 120
QY 404 CACGACGCTCTCTTACAGGGAAGCGGCTTTTAAAGATGCTCAGCAATAGAGAGTCTGT 463
DB 121 HisSerSerSerTyrArgGlyAlaValProLysMetLeuSerAsnLysGlnSerLeu 140
QY 464 GATCAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 523
DB 141 AsnGlnIleLeuValAlaTrpHisGlyGlnLysValLeuValLeuValValValVal 160
QY 524 CAGGACAGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 583
DB 161 GlnAspLysAspLeuPheSerValPheAlaAspProAsnMetLeuAspSerLeuValPro 180
QY 584 GTCACCCAGCCCTCGTCAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 643
DB 181 AlaHisProAlaLeuValAsnAlaLeuValLeuValLeuHisSerValAlaGlySerAla 200
QY 644 CCAATGCTGGGAGCTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
DB 201 ProMetProGlyThrAspSerSerArgSerMetProSerSerSerTyrArgAspSer 220
QY 704 CCAAGTGGCTTCTGTTGAGGCTGCTCAGATGATGATGATGATGATGATGATGATGAT 763
DB 221 ProGlyGlyPheLeuPheGluGlyLeuSerAspAspGlnAspPheHisProAsnThr 240
QY 764 AGTTCACACCCCTTAGCAGTACCCAGCTCCCGCCGACCCCTCTCTCTCTCTCTCTCT 823
DB 241 ArgSerThrProSerSerThrProSerArgProAlaSerLeuGlyTyrSerGly 260
QY 824 GCTGCTGGGCGCCCGCCCATCACCCAGAGTGGCTGGCCACCGCTTGGCCCTGGCCAG 883
DB 261 AlaAlaGlyProArgProIleThrGlnSerGlnLeuAlaThrAlaLeuAlaLeuAlaSer 280

```

```

QY 884 ACTCCGAGAGCAGCTCTCACACACCGACTCTCTGGCACCAGGCTCATTTCTCAGGACC 943
DB 281 ThrProGlnSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
QY 944 TCACCAATGCTCTCTGGTGTCCAGTCAGGACGCCCATCCCAATGATCTCTTTCAGCCAA 1003
DB 301 SerProMetSerSerGlyValGlnSerGlyThrProIleThrAsnAspLeuPheSerGln 320
QY 1004 GCCCTACAGCATGCCCTTCAGGCTCTGGGACGCCCATCCAGCCAGTGGCAGCCC 1063
DB 321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnTrpGlnPro 340
QY 1064 CAGCTGCAGCAGCTACGCTGACATCGGCATCCAGGACCATCAGCTGAGCCTGGGCCCCTGC 1123
DB 341 GlnLeuGlnGlnLeuAlaGaspMetGlyLeuGlnAspAspGluLeuSerLeuArgProCys 360
QY 1124 AGGCATCGTGGGACATCCACAGCCCTCGGAGCTCATCTTTCTGGAGGAGCCCAT 1183
DB 361 ArgProProValGlyThrSerLysGlnProTrpSerSerSerLeuLeuGluProHis 380
QY 1184 GAATCTCTGCTCTCCCTGAACCCCGACCAAGTTGACAGGCTACTGCGCTTGGGAGGCA 1243
DB 381 GlnLeuProAlaSerProGluProProAlaSerCysArgGlyTyrCysProTrpGluAla 400
QY 1244 CTCATGAGGTGCCCTCCATCTCTCCCTGTC 1273
DB 401 LeuMetLysValProProSerLeuProVal 410

```

RESULT 2

US-08-725-758A-2
 Sequence 2, Application US/08725758A
 Patent No. 61601c8
 GENERAL INFORMATION:
 APPLICANT: Reed, Guy
 APPLICANT: Clerent, Christophe Y.
 TITLE OF INVENTION: NOVEL PLATELET ACTIVATION PROTEIN
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Fish & Richardson P.C.
 STREET: 225 Franklin Street
 CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02110-2824
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FASTSEQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/725,758A
 FILING DATE: 04-OCT-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 60/005,074
 FILING DATE: 06-OCT-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Fraser, Janis K.
 REGISTRATION NUMBER: 34,819
 REFERENCE/DOCKET NUMBER: 05433/020001
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 617-542-5070
 TELEFAX: 617-542-8906
 TELEX: 200154
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 426 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 FRAGMENT TYPE: internal
 US-08-725-758A-2

```
Alignment Scores:
Pred. No.: 3,29e-182 Length: 426
Score: 2129.00 Matches: 429
Percent Similarity: 100.00% Conservative: C
Best Local Similarity: 100.00% Mismatches: C
Query Match: 92.48% Indels: C
DB: 3 Gaps: C

US-09-735-251-3 (1-1273) x US-09-735-251-3 (1-426)

Qy 44 ATGTCCTCTCAGACTGGCAGCTTCGGTGAAGTGGCTGACCCAGGCGCTTACCTCCAAAG 103
Db 1 MetSerLeuSerAspTrpHisLeuAlaValLysLeuAlaAspGlnProLeuThrProLys 20
Qy 104 TCTATTCCTGGTGGCAGACAGACTGGAGAAATACCTGCTAGCGGCTATAGTATT 163
Db 21 SerileLeuArgLeuProGlnThrIleSerLeuGlyIleTySerLeuGlyGlyTySerIle 40
Qy 164 TCATTTCGAGGAGCTTATTGCTGCGAAAGTCCAGGAGTGTTCAGAGCCCTGAGCTG 223
Db 41 SerPheLeuLysGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 50
Qy 224 ATTGATCTGATCTACTGTGGTGGAGGTTAAAGATGACCCAGACACTTGTCTTATGG 283
Db 61 IleAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
Qy 284 ATTCAACTGGGTGCTACTGTCCATTTCTGGGAAAGTCTGCTGAACTGAACTGATCAAAA 343
Db 81 IleGlnProGlySerThrValHisValLeuArgLysSerTrpProGluProLys 100
Qy 144 CCGAACTGTGGGACAAAGTGGTGGCAGTGGAGAGTTCGGGGTGTGGCACTGCGCTG 403
Db 101 ProGlnProValAspLysValAlaValSerArgLysLeuLeuLeuLeuLeuLeuLeu 120
Qy 404 CACAGCACTCTCTTACAGGAGTGGTCTTTTAAAGATGCTCAGCAATAAGAGTCTCTG 463
Db 121 HisSerSerSerTyArgGluAlaValPheLysMetLeuSerAsnLysGluSerLeu 140
Qy 464 GATCAGATCATTTGGCCACCCCAAGCTTCAGTATGACCTATGCTTCTTGGGGTCTG 523
Db 141 AspGlnIleValAlaThrProLysLeuSerSerAspProIleAlaLeuGlyValLeu 160
Qy 524 CAGCAGAGGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 583
Db 161 GlnAspLysAspLeuPheSerValIleAlaAspProAsnMetLeuAspThrLeuValPro 180
Qy 584 GCTCACCAGCCCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 643
Db 191 AlaHisProAlaLeuValAspAlaValValIleValLeuHisValValAlaGlySerAla 200
Qy 644 CCAATGCTGGATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 703
Db 201 ProMetProGlyThrAspSerSerArgSerMetProSerSerSerTyArgAspMet 220
Qy 704 CCAGTGGCTCTCTGTTGAAGGCTCTCAGACATGAGGATGATTCACCCAAACAC 763
Db 221 ProGlyGlyPheLeuPheGluGlyLysSerAspAspGluAspPheHisProAsnThr 240
Qy 764 AGGTCCACACCTCTAGCAGTACTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
Db 241 ArgSerThrProSerSerSerThrProSerSerArgProAlaLeuLeuGlyTySerGly 260
Qy 824 GCTGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 883
Db 261 AlaAlaGlyProArgProIleThrLysSerGluLeuAlaThrAlaLeuAlaLeuAlaSer 280
Qy 884 ACTCGGAGAGCAGCTCTCAGACAAGCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 943
Db 281 ThrProGluSerSerSerHisThrProThrProGlyThrGlnGlyHisSerSerGlyThr 300
Qy 944 TCACCAATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1003
Db 301 SerProMetSerSerGlyValGlnLysLeuThrIleThrAspAspLeuPheSerGln 320
```

```

1004 GCCTTACAGATGCGCTCTTCCAGGCTCTGGGAGCCGCCAGCTTTCAGAGCCAGTGGGAGCCC 1063
321 AlaLeuGlnHisAlaLeuGlnAlaSerGlyGlnProSerLeuGlnSerGlnTrpGlnPro 340
1064 CAGCTCCAGCAGCTACGTCACATGGCATCCAGGAGGATGAGCTGAGCTGGGCGCTGCG 1123
341 GlnLeuGlnGlnLeuArgAspMetGlyIleGlnAspGluLeuSerLeuArgProCys 360
1124 AGCCACCCGGTGGGACATCCAGCAGCCCTGGAGCTCATCTTTTGTGAGGAGCCCAT 1183
361 ArgProProValGlyThrSerLysGlnProTrpSerSerSerLeuLeuGluGluProHis 380
1184 GAATCTCTCTGCTTCCCTGAACCCCGAGCAAGTTCAGAGGCTACTGCTTGGGAGGCA 1243
381 GlnLeuProAlaSerProGluProProAlaSerCysArgGlyTyTrpCysProTrpGluAla 400
1244 CTCATCAAGGTGCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1270
401 LeuMetLysValProProSerLeuPro 409

RESULT 3
US-09-266-225D-12
; Sequence 12, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernev, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 816
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-266-225D-12

Alignment Scores:
Pred. No.: 3,64e-06 Length: 816
Score: 163.00 Matches: 88
Percent Similarity: 36.48% Conservative: 28
Best Local Similarity: 27.67% Mismatches: 125
Query Match: 7.08% Indels: 77
DB: 4 Gaps: 11

US-09-735-251-3 (1-1273) x US-09-266-225D-12 (1-816)

Qy 311 CTGCGAAAGTCTGGGCTCGAACCTGATCAGAAACCGGAACCTGTGCAAAAGTGGCTGCC 370
Db 491 LeuArgAspGlyProSerAlaProLeuGluAlaProGluProArgLysProValThrAla 510
Qy 371 -----ATGAGAGATTCCGGGTGTTGCCACTGCTCCCTGCACAGCAGCTCTCTTAC 421
Db 511 GlnGluArgSerLysArgGluGluLysArgArgArgGlnGluArgAlaLysGlu 530
Qy 422 AGGAGGCGCTCTTTAAGATGCTCAGCAATAGGAGTCTCTGATCATGATCATTTGTGGCC 481
Db 531 ArgGluLys-----LysArgArgGlnGluArgGluArgLysGluArgGlyAlaGlyAla 547
Qy 482 ACCCCAGGCTCAGCAGTCACTTGTCTTGGGTTCTCCAGCAAGAGCCTCTTTC 541
Db 548 SerGlyGlyProSerThrAspProLeuAlaGlyLeuValLeuSerAspAsnArgSer 567
Qy 542 TCTGCTCTTCTGCTATCCCAATATGCTTGTATGATGTTGCTGCTGCTGCTGCTGCTGCTGCT 601
Db 567 rLeu-----LeuGluArgTrpThrArgMetAlaArgProAlaAla 580
```

[illegible]

RESULT 4

JS-CG-579 1A:-2

; Sequence 2, Application US/9507338X;

: Patient No. 6365372

GENERAL INFORMATION:

APPLICANT: Christina Jones

APPLICANT: Yacik, Peter

TITLE OF INVENTION: SHEET 3

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

ADDITIONAL INFORMATION REQUESTED

CURRENT EXPIRATION DATE: 2000

CURRENT FILING DATE: 2003-03-03
PRICE PER SHARE: 10.00

; PRIOR APPLICATION NUMBER: 1000 00
 PRIOR FILING DATE 1000 00

; PRIOR FILING DATE: 1999-05-17

; NUMBER OF SEQ ID NOS: 1 /

; SOFTWARE: Patent in Ver. 2.

: SEQ ID NO 2

; LENGTH: 2972

```

: type: prt

```

; ORGANISM: Human

JS-09-579-181-2

Alignment Scores:

Pred. No.: 0.0002

Score: 145.50

Percent Similarity: 33.80%

Best local similarity: 23.74%

Query: Match.

Quesada, M. A. 1970. A

22

```

QY 990 ----ATCTCTCAGCCAGCCCTACAGCATCCCTTC----- 1022
Db 1181 rSerLeuAlaSerAlaSerProValProAlaProThrProValLeuAlaProSerSerTh 1401
QY 1023 ----AGCTTCTGGGAGAGCCAGCTTCAGAGCCA- 1053
Db 1401 rGlnThrMetLeuProAlaProValProSerProLeuProSerProAlaSerThGlnT 1421
QY 1054 ----GTGGCAGCCAGCCAGCTCAGAGCTAGCTAGCATGGG CATCCAGAGAGATGAG 1105
Db 1421 hrLeuAlaLeuAlaProAlaLeuAlaProThrLeuGlyGlySerSerProSerGlnThr 1440
QY 1106 CTGAGCCCTGGCGGCTCAGAGCCA-----CGGTGGGAGAGATCAGAGCCG 1153
Db 1441 LeuSerLeuGlyThrGlyAsnProGlnGlyProGlnProGlnThrLeuSerLeuThr 1460
QY 1154 TGGAGCTCATCTTCTCTGGAGAGCCCATGMA----- 1186
Db 1461 ProAlaSerSerLeuValProThrProAlaGlnThrLeuSerLeuAlaProGlyProPro 1480
QY 1187 ----CTCCCTCTCTCCCTGAGCCAGCCAGCAAGTGC 1219
Db 1481 LeuGlyProThrGlnThrLeuSerLeuAlaProAlaProProLeuAlaProAlaSerPro 1500
QY 1220 AGAGGCTACTCCCTGGAGAGCACTGTAAGTGGCTGATCT 1264
Db 1501 ValGlyProAlaProAlaHisThrLeuThrLeuAlaProAlaSer 1515

RESULT 5
US 09-579-181-1
; Sequence 1, Application US/09579-81
; GENERAL INFORMATION:
; APPLICANT: Chiviva, John
; APPLICANT: Yachuk, Peter
; TITLE OF INVENTION: SNF2 Related GTP Activator Protein (SNCAP)
; FILE REFERENCE: 6153-4247
; CURRENT APPLICATION NUMBER: US/09/579-181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1

Alignment Scores:
Pred. No.: 0.00242 GenPos: 100%
Score: 145.50 Matches: 114
Percent Similarity: 31.80% Conservative: 50
Best Local Similarity: 23.74% Mismatches: 141
Query Match: 6.32% Indels: 143
DB: 4 Gaps: 2

US-09-735-251-3 (1-1273) x US-09-579-181-1 (1-3118)
QY 77 CTGGGTGACCCAGCCACTTACTCCAAAGCTATATT----- 112
Db 1214 LeuAlaProAlaProArgProProArgSerGlyLeuProAlaValLeuAsnProArgPro 1233
QY 113 -----CGGTGCCAGAGACAAATGGGAGAAATACCTCGCTAGGGGGTAT 157
Db 1234 ThrLeuThrProGlyArgLeuProThrProThrLeuGlyThrAlaAlaProMetPro 1253
QY 158 AGTATTTCATTCTGAGCAGCTTATCTGGAAACTCCAGAGACTCGTTCCAGAGCCT 217
Db 1254 ThrProThrLeuValArgProLeuLeu-----LysLeuValHisSerProSerPro 1270
QY 218 GAGCTGATTGATCTGATCTACTGTGTGGAGAGTAAAGATGACACAGACATTGACTTC 277

```

```

Db 1271 GluVal-----SerAlaSerAla 1276
QY 278 TATGGCATTCACCT-----GGGTCCACTGTCCATGTTCTCGAAAGTCTGGGCT 328
Db 1277 ProGlyAlaAlaProLeuThrIleSerSerProLeuHisVal---ProSerSerLeuPro 1295
QY 329 GAACCTGATCAGAAACCGGAACCTGTGGACAAAGTGGCTGCCATGAGAGAGTTCGGGTG 388
Db 1296 GlyProAlaSerSerProMetProIleProAsnSerSerProLeu----- 1310
QY 389 TTGCACACTGCCCTGCACAGCAGCTCTCTTACAGGAGGGGGTCTTTAAGATGCTCAGC 448
Db 1311 -----AlaSerProValSerSerThrValSerValProLeuSer 1323
QY 449 AATAAGAGTCTCTGATCAGATCATTTGGCCACCCAGGAGCTCAGCAGCTCAGCCTATT 508
Db 1324 SerSerLeuProIle---SerValProThrThrLeuProAlaProAlaSerAlaProLeu 1342
QY 509 GCTCTTGGGGTCTCCAGGACAGGACCTCTCTCTGCTGCTGCTGATCCCAATATGCTT 568
Db 1343 ThrIleProIleSerAlaProLeu----- 1350
QY 569 GATACCTTGGTCTGCTGCACCCAGCCCTCGTCAATGCAATTGCTCTGCTGTCACATCC 628
Db 1351 ---ThrValSerAlaSerGlyProAlaLeuLeuThrSerValThrProProLeuAlaPro 1369
QY 629 GAGCAGGAGTGCCCCAATGCTGGAGTACTCTCTTCCCGAGGATGCCCTCCAGC 688
Db 1370 ValValProAlaAlaProGlyProProSerLeuGlnProSerGlyAlaSerProSerAla 1389
QY 689 TCATACCGGGATATGCCAGGTGGCTTCTGTTGAAGGGCTCTCAGATGATGAGATGAC 748
Db 1390 SerAlaLeuThrLeu---GlyLeuAlaThrAlaProSerLeuSerSerSerGlnThrPro 1408
QY 749 TTTCACCA-----AACACACAGGTCC 769
Db 1409 GlyHisProLeuLeuLeuAlaProThrSerSerHisValProGlyLeuAsnSerThrVal 1428
QY 770 ACACCTCTTAGCAGT-----ACTCCAGCTCCCGCCAGCTCCCTCGGGTACAGT 820
Db 1429 AlaProAlaCysSerProValLeuValProAlaSerAlaLeuAlaSer---ProPhePro 1447
QY 821 GGAGCTGTGGCGCCGCCCATCACCAGAT----- 853
Db 1448 SerAlaProAsnProAlaProAlaGlnAlaSerLeuLeuAlaProAlaSerSerAlaSer 1467
QY 854 ---GAGTGGCCACCCGCTTGGCCCTGGCCAGCAGCTCCGAGAGCAGCTCTCACACCG 910
Db 1468 GlnAlaLeuAlaThrProLeuAlaProMetAlaAlaProGlnThrAlaIleLeuAlaPro 1487
QY 911 ACTCTGTGACCCAGGGTCACTCT-----CAGGGACCTCACCAGTGTCC 955
Db 1488 SerProAlaProProLeuAlaProLeuProValLeuAlaProSerProGlyAlaAlaPro 1507
QY 956 TCTGTGTCCAGTCAGGAGCGCCCATCACCAGT----- 989
Db 1507 oValLeuAlaSerSerGlnThrProValProValMetAlaProSerSerThrProGlyTh 1527
QY 990 ----ATCTCTTCCAGCAAGCCCTACAGCATCCCTTC----- 1022
Db 1527 rSerLeuAlaSerAlaSerProValProAlaProThrProValLeuAlaProSerSerTh 1547
QY 1023 -----AGGCTCTGGGAGCCAGCCAGCTTCAGAGCCA- 1053
Db 1547 rGlnThrMetLeuProAlaProValProSerProLeuProSerProAlaSerThrGlnT 1567
QY 1054 -----GTGGCAGCCAGCTCAGGAGCTACCTGACATGGG---CATCCAGGACCATGAG 1105
Db 1567 hrLeuAlaLeuAlaProAlaLeuAlaProThrLeuGlyGlySerSerProSerGlnThr 1586
QY 1106 CTGAGCTCGGGCCCTGCAGGCCA-----CCGTTGGGAGCATCCCAAGCAGGCC 1153

```


Score: 142.50 Matches: 105
 Percent Similarity: 34.23% Conservative: 47
 Best Local Similarity: 23.65% Mismatches: 174
 Query Match: 6.19% Indels: 118
 DB: 4 Gaps: 16

US 09-735-251-3 (1273) x US-09-252 901A-21827 (1-1093)

QY 92 CTTATCGAAGCTCTATTCTGGTTCGAGTCCAGAACACAGACATGAGAGATATGCTAGGG 151
 DB 441 ValThrProValSerLeuThrArgThrProValProArgThrPheArgThrProVal 457
 QY 152 GGCATAGTATTTCATTCGAGGAGATTTATGCTGGCAGACACAGAGCTCTGTTCGA 211
 DB 458 -----PheArgLeuAlaSerSerArgThrProValLeuThrArgThrProVal 474
 QY 212 GACCTGAGCTGATTCGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCT 271
 DB 474 o-----LeuSerSerArgAlaAlaThrArgValMe 484
 QY 272 GACTTCTATGGATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 323
 DB 484 tProAlaLeu-----LeuProMetProProProLeuArgLeuLeuSerSerThrProValArg 503
 QY 324 -----GGCTGGAACCTGATCAGAACCGGACCTGTGGACAAAGTGGCTGCCATG 373
 DB 503 eSerThrAlaProLeuAlaLeuThrProArgLeuLeuSerArgLeuAlaPro 522
 QY 374 AGAGAGTTCGGGTTTCACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 433
 DB 523 -----CysSerValThrProAlaSerLeuSerSerLeuProProThrPLeuSe 539
 QY 414 -----GCTCTTAC 421
 DB 538 rSerAlaAlaTrpLeuValSerAlaSerAlaProAlaLeuGluGlyValProProCArg 558
 QY 422 AGGAGGGGGTCTTAAAGATCTCAAAATAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 466
 DB 558 uSerArgAlaAlaAlaAlaAlaValArgLeuProSerLeuThrSerAlaLeuProThrLe 578
 QY 467 CAGATCATTTGGCCACCCAGGCTTACAGATCAGCTTATTTCTGCTGCTGCTGCTGCT 526
 DB 578 uSerSerThr-----ProProArg-----LeuThrLeuAlaLeuThrPLeuSerSe 594
 QY 527 GACAGAGACCTCTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 586
 DB 594 rArgProProSerProLeuAsnSerAlaProProSerArgLeu-----ArgSe 610
 QY 587 CAGCAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 637
 DB 610 oSerArgProAlaSerThrProLeuAlaThrProSerArgAlaAlaThrAlaArgAla 630
 QY 638 AGTSCCCCAATGCCCTGGACCTGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 697
 DB 630 gProLeuSerProMetThrLeuProProArgLeuSerSerCysSerArgAlaPheThrAl 650
 QY 638 GATATGCCAGCTG-----GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
 DB 650 aThrPheGluValLeuGluLeuSerProAlaAlaLeuSerThrCysArgAlaSerIleAl 670
 QY 734 GATGATGAGGATG----- 746
 DB 670 aMetProProPheAlaAlaIleSerGlnAspTrpLeuLeuSerIleValSerAlaAlaThr 590
 QY 747 -----ACTTTCCACCAACACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 799
 DB 690 rPheArgValPheSerLeuThrSerSerProProCysTrpAlaArgLeuProSerValAl 710
 QY 803 C-----CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 835
 DB 710 aCysArgSerProLeuAlaAlaIleArgProProAlaLeuProThrAlaPheAlaAlaAr 730
 QY 836 CGGCCCATCACCAGAGTGAAGTGGCCACCGCTTGGCCCTGGCCAGCACTCCGGAGAGC 895

DB 730 gSerArgArgProSerLeuAsnSerLeuProProTrp-----SerLeuArgArgLe 747
 QY 896 AGCTCTCACACACCGA-----CTCCTGGCACCAGGGTCATTCTCTCA 937
 DB 747 uAlaThrSerThrArgThrProAlaLeuProLeuAlaLeuProProSerArgLeuSerLy 767
 QY 938 GGGAGCTTACCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
 DB 767 sAlaProProValAlaLeuLysProAlaSerAlaThrSerThrProProArgLeuSerAr 787
 QY 995 AGCCAAGTCTACAGCATGCCCTTCAGGCTCTGGCCAGC---CCAGCTTCAGAGCCAG 1054
 DB 787 gluAlaArgThrThrPheAsnSerLeuArgLeuAspSerArgProProAlaArgLeuSe 807
 QY 1055 TGGCAGCTCCACCTGACAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1114
 DB 807 rSerCysProThrValThrArgCysAlaProTrpLeuSerArgLeuProSer--SerArg 826
 QY 1115 CGGCCTTCAGCCACCGCTGGGACATCCAGAGCCCTGAGCTCATCTTCTGCTGAG 1174
 DB 827 LeuSerSerArgProProArgLeuThrSerArgProPro-----ArgLeuLeuSer 843
 QY 1175 GAGCCC 1180
 DB 844 ValPro 845

RESULT 8

US-08-396-479B-6
 ; Sequence 6, Application US/08396479B
 ; Patent No. 5612455
 ; GENERAL INFORMATION:
 ; APPLICANT: HOBY, Timothy
 ; TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR, HORNBACH, TEST, ALBRITTON & HERBERT
 ; STREET: 4 Ebarcadere Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/396.479B
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Osman, Richard A
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: A-59450-1/RAO
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 494-8700
 ; TELEFAX: (415) 494-8771
 ; TELEX: 210 277299
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 902 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-396-479B-6

Alignment Scores:
 Pred. No.: 0.000288 Length: 902
 Score: 142.00 Matches: 91
 Percent Similarity: 35.71% Conservative: 29
 Best Local Similarity: 27.08% Mismatches: 102
 Query Match: 6.17% Indels: 115

```

58:                                     1           Gaps:                20
US-09-735-251-3 (1-1273) x US-08-818-823-6 (1-902)
QY 472 CATTGTGGCCACCCAGGCTTACGCTAGCTTATTGCTCTGG          516
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 HissrProProArgProAlaProSerProGlyThrProGluSerGlnProAla 90
QY 517 -----GGTCTCCAGACAGAGACT----- 537
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 ArgSerValArgLeuGlyProGlyGlyGlyAlaGlyGlyGlyArgVal 110
QY 538 CTTCTCTGT-----CTTCGCTGATCCCAATATGCTTTGATCGCTGCTCACC 591
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 LeuGlyCysProSerIleArg-----IleThrSerIleSerPro 123
QY 592 AGCCCTCGTCAATGCCATTGTCTCTCTTGTACTCCGTAGCAGCAGTGCCTCAATGCC 651
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 ThrProGlu-----ProProAlaAlaLeuGluAspAsn-----ProAspAla 137
QY 652 TGGGACTGACTCTCTCCCGAGCATCGCTTC-----AGCTCATACCGGA 699
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 TrpGly-AspGlySerProArgAspTyrProProGluGlyPheGlyTyrArgI 157
QY 700 T-----ATGCCAGTGGCTCTCTTTTGAAGAGTCTCAGATGATGATGACTTTCA 753
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 AlaGlyAlaGlnGlyGlyAlaPhe-----Phe 168
QY 754 CCCAACACACAGTCCACACCCCTTACCAAT-----ACTGAGCTCCGCGC 801
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
168 rProSerProGlySerSerSulLeuSerIlePhePheSerIleAlaSerAsp 189
QY 802 AGCTCCCTCGGGGTACAGT----- 820
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
188 AlaAlaLeu-----TyrAlaAlaGly-AspGluIleLeuSerValLeuAsnAlaAla 207
QY 821 -----GGAGTGTCTGGCCCGGCTGATATATATAGAGTAA 264
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
207 rArgPheGlyLeuGlySerProLeuProSerProArgAlaSerProArgPro 227
QY 865 CGCTTGGCGCTGGCCAGCACTCGAAGACAGTCTCACAACGATCTCGGACCA 924
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
227 rGluAspProTrp-----SerLeuGlyGlyPro 237
QY 925 GGGTATTATTCAGGAGGCTAAATGCTGCTTTTATATATATATATATAT 984
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
237 rPheGlyGlyArgGlyAlaGlnLeuGlnLeuGlnLeuGlnLeuGlnLeu 257
QY 985 CAATGCTCTCTTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1044
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
257 rGlu-----AlaSerProArgProAlaSer-----ProGlyGlyLysArgArg 271
QY 1045 TCAGACCCAGTGGCAGCCCGAGCTGAGCAG-----CTAGCTGACATGGCATCCAG 1097
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
271 rYserSerSerGlyThrProSerIleAlaSerProAlaLeuSerArgArgGly 291
QY 1098 ACCATGAGCTGAGCTCGCCGCTGAGGCTATGCGTGGGAGATCCAGCAGCCTG 1157
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
291 rGluGluGlySerGluProProLeuProLeuProLeuProLeuProLeu 311
QY 1158 GC----- 1159
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
311 rPheGlyProPheAspTyrValAlaProProAlaGluSerIleProGlnLysThr 331
QY 1160 --TCATCTTTGTGGAGGAGCCCATGAGTCTGCTCTGCTGCTGCTGCTG 1217
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
331 rArgThrSerSerGluGlnAlaValAlaLeuProArgSerGluGluProAlaSer 350
QY 1218 GCAGAGGCTACTGCCCTTGGAGGAGTCAATCATTAAGTGGCTCA 1261
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
350 rAsnGlyLysLeuProLeuGlyAlaGluGluSerValAlaPro 364

```

RESULT 9

```

US-08-818-823-6
: Sequence 6, Application US/08818823
: Patent No. 5708158
: GENERAL INFORMATION:
: APPLICANT: HOEY, Timothy
: TITLE OF INVENTION: NUCLEAR FACTORS AND BINDING ASSAY
: NUMBER OF SEQUENCES: 18
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
: STREET: 4 Embarradero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/818,823
: FILING DATE: 14-MAR-1997
: CLASSIFICATION: 536
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/396,479
: FILING DATE: 02-MAR-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-59450-1/RAO
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELETYPE: 210 277299
: INFORMATION FOR SEQ ID NO: 6:
: LENGTH: 902 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: Protein
US-08-818-823-6

Alignment Scores:
Pred. NO.: 6,000288 Length: 902
Score: 142.00 Matches: 91
Percent Similarity: 39.71% Conservat: 29
Best Local Similarity: 27.08% Mismatches: 102
Query Match: 6.17% Indels: 115
D5: 1 Gaps: 20

US-09-735-251-3 (1-1273) x US-08-818-823-6 (1-902)
QY 472 CATTGTGGCCACCCAGGCTTACGCTAGCTTATTGCTCTGG----- 516
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 HissrProProArgProAlaProSerProGlyThrProGluSerGlnProAla 90
QY 517 -----GGTCTTCAGGACAGAGACT----- 537
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
91 ArgSerValArgLeuGlyGlyProGlyGlyGlyAlaGlyGlyGlyArgVal 110
QY 538 CTTCTCTGT-----CTTCGCTGATCCCAATATGCTTTGATCGCTGCTCACC 591
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
111 LeuGlyCysProSerIleArg-----IleThrSerIleSerPro 123
QY 592 AGCCCTCGTCAATGCCATTGTCTCTCTTGTACTCCGTAGCAGCAGTGCCTCAATGCC 651
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
124 ThrProGlu-----ProProAlaAlaLeuGluAspAsn-----ProAspAla 137
QY 652 TGGGACTGACTCTCTCCCGAGCATCGCTTC-----AGCTCATACCGGA 699
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
138 TrpGly-AspGlySerProArgAspTyrProProGluGlyPheGlyTyrArgI 157
QY 700 T-----ATGCCAGTGGCTCTCTTTTGAAGAGTCTCAGATGATGATGACTTTCA 753
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```


DQ 590 ArgLeuSerArgSerProPheThr:SerThrLeuProSerSerSerAlaSerIleSerThr 609
 QY 125 ACAGAACTGGGAGAACTACGCTAGGGGCTATAGTATTCTCTGAGCAGCTTAT 184
 Db |||||
 DQ 610 ThrGlnVal:-----PheSerLeuAlaGlySer:ThrPheSerLeuProSerThrHisIle 627
 QY 185 GCTGGC-----AAACTCCAGGAGTCTGTCTCA 211
 Db |||||
 DQ 628 PheGlyThrProMetGlyAlaValAsnProLeuLeu:SerGlnAlaGluSerSerHisThr 647
 QY 212 GACCTGAGCTGATTGATCTGATCTGCTGGGAGAACTAAAGATGACGAGCACTT 271
 Db |||||
 DQ 648 GlnProAspLeuGluAspCysSer:PheArgCysArgGlyThrSerProGlnGluSerLeu 667
 QY 272 GACTTCTATGGCATTCACTGCTGCTCACTGTCATGTT 310
 Db |||||
 DQ 668 Ser:-----SerValSerProIleSerSerLeuProAlaGluPheAspGlnThrAlaSer 685
 QY 311 -----CTGCGAAAGTCTGSCCTGAACCTGATGAGAAACCGGAA 349
 Db |||||
 DQ 686 AlaProCysGlyGlyGlnLeuAspProAlaAlaProGlyThrThrAsnValGluGln 705
 QY 350 CCTGTGGACAA-----GTGCTGCTATGATGAGAGAGTCTCCGCTG 388
 Db |||||
 DQ 706 LeuLeuGluLysGlnGlyAspGlyGlnAlaGlyValAsnIleValGluMetLeuLysAla 725
 QY 389 TTCACACTGGCCCTGCACAGCAGTCTCTCTTAAGGAGGCTCTTTAAGATGCTCAGC 448
 Db |||||
 DQ 726 LeuHisAlaLeuGlnLysGlnAsnGlnArgGluGlnGluGlnIleLeuSerLeuThrAla 745
 QY 449 AATAGAGCTCTCTGATCAGATCATATGTGGACCCCGACCCGAGGCTGAGCACTAT 508
 Db |||||
 DQ 746 LysLysGluArgLeuGlnIleLeuAsnVal----- 755
 QY 509 GCTCTTGGGCTCTCTACAGACAGACCTCTCTCTGCTCTGCTCTGCTGATCCCATATGTT 568
 Db -----GlnGlu 757
 QY 569 GATACGTTGGTCTGCTCAGCTGAGCTGCTGGTAATGCCATGCTGCTGCTGCTGCTGCTG 628
 Db |||||
 DQ 758 SerValProPheProAlaLeuPheAlaAlaProAlaAlaAlaGlyProAlaProGly 777
 QY 629 GTACAGGAGTCTCCCAATGCCGAGGACTGAG-----TCCCTCTCCCGGAGCAG 679
 Db |||||
 DQ 778 ProTyrGlyLeuProProGlnAlaLysSerSerAspSerLeuSerThrSerLysSerPro 797
 QY 680 CCTCCAGCTATACGCTGATATATAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 739
 Db |||||
 DQ 798 ProLysSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 812
 QY 740 GAGTATGACTTTTCACTGAAAT 799
 Db |||||
 DQ 813 SerGluAsp-----ProHisSerGlyCysProSerArgSer 824
 QY 800 CCAGCTCTCTCTGGGTACAGTGAATGTGCTGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTG 859
 Db |||||
 DQ 925 SerSerSerLeuSerPheHisSerThrProPheProLeuPheLeuLeuLeuLeuLeuLeuLeu 844
 QY 860 GGCACC-----CCCTTGGCCCTGGTACACACTCTGAGAGAGAGAGAGAGAGAGAGAGAG 910
 Db |||||
 DQ 845 AlaThrLeuProLeuAlaLeuPheGlyAlaProAlaProAlaProAlaProAlaProAlaPro 864
 QY 911 -----ACTCTGTGCAGGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 958
 Db |||||
 DQ 865 GlyLeuGlyArgAlaProGlyAlaAlaGly-----LeuGlyAlaValProMetAlaGlu 882
 QY 959 GGTCTCCAGTACGAGGAGCCCATATACCAATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1018
 Db |||||
 DQ 883 GlyLeuLeuGlyGlyLeuAlaGlySerGlyLysLeuProLeuAsnGlyLeuLeuGlyGly 902
 QY 1019 CTT-----CAGGCCCTCTGGCAGCCAGC 1042
 Db |||||
 DQ 903 LeuAsnGlyAlaAlaAlaProAsnProAlaSerLeuSerGlnAlaGlyAlaProThr 922

QY 1043 CTTTCAG-----AGCCAG 1054
 Db |||||
 DQ 923 LeuGlnLeuProGlyCysLeuAsnSerLeuThrClnGlnArgHisLeuLeuGlnGln 942
 QY 1055 TGGCAGCCCTAGCTGCGAGCAGCTACGTGACATGGGCACTCCAGGACGATGAGCTGAGCCTG 1114
 Db |||||
 DQ 943 GlnGlnGlnGlnLeuGlnGlnGln-----GlnLeuLeuAla 955
 QY 1115 CGGCCCTGCAGCCACCG----- 1132
 Db |||||
 DQ 956 SerProGlnLeuThrProGluHisGlnThrValValTyrGlnMetIleGlnGlnGln 975
 QY 1133 -----GTGGGACATCCCAAGCAGCCCTGAGGCTCA 1162
 Db |||||
 DQ 976 GlnLysArgGluLeuGlnArgLeuGlnMetAlaGlyGlySerGlnLeuProMetAlaSer 995
 QY 1163 TCTTGTCTGAGGAG-----CCCATGAATCCCTGCT 1195
 Db |||||
 DQ 996 LeuLeuAlaGlySerSerThrProLeuLeuSerAlaGlyThrProGlyLeuLeuProThr 1015
 QY 1196 TCCCTGAAACCCAGCAAGTTGCGAGAGGCTACTGCCCTTGGGAGGCACTCATGAAGGTG 1255
 Db |||||
 DQ 1016 AlaSerAlaProPro-----LeuLeuProAlaGlyAlaLeu-----Val 1028
 QY 1256 CTTCCATCTCTC 1267
 Db |||||
 DQ 1029 AlaProSerLeu 1032

RESULT 13

US-09-252-991A-30843
 ; Sequence 30843, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 127196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30843
 ; LENGTH: 663
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-30843

Alignment Scores:

Pred. No.: 0.000865 Length: 663
 Score: 136.00 Matches: 104
 Percent Similarity: 29.42% Conservativeness: 29
 Best Local Similarity: 23.01% Mismatches: 147
 Query Match: 5.91% Indels: 172
 DB: 4 Gaps: 22

US-09-735-251-3 (1-1273) x US-09-252-991A-30843 (1-663)

QY 286 TCACCTGGTCCACTGTCCATGTCTGCGAAAGTCTCTGGCCTGAACCTGATCAGAACC 345
 Db |||||
 DQ 209 SerAlaTrpLysProCysAlaCys---AlaLysAlaValGlySerAlaArgSerProAla 227
 QY 346 GGAACCTGTGCACAAAGTGGCTGCCATGAGAGAGTTCGGGTGTGTGCACACTGCCCTGCA 405
 Db |||||
 DQ 228 --AsnAlaTrp-----ProAlaAlaSerAlaTrpCysProThr-----CysA 241
 QY 456 CAGCAGCT-----CCTCTTACAGGAGGC 429
 Db |||||
 DQ 241 snSerAlaArgProSerAlaIleAlaSerThrAlaArgArgSerProSerThrAsnGlyA 261


```
Ds 134 ..... GlyArgThrArgProGlyArgGluArg 143
Qy 728 GCCTTCAACAGGAGCAC...GTGGATATCCGGATGAGCTGGAGGCGATGCTCC 672
Db 144 CysLeuArgAlaProAlaArgProValAlaGlyProProAlaGlyTyrPArgGluLeuPro 163
Qy 671 GGGAGAGAGAGTCACTCCAGGAGTGGAGCA...TTCCTGGTATGAGTSCAGACACGAGA 612
Db 164 ProAlaGlySer...AlaArgAlaArgGly...ArgAlaProGluProGly 178
Qy 611 CAATGGCATTCAGAGGGTGGH...GAGCAGGAGCAACAGATTAAGCATATGGGATCAG 552
Db 179 Arg.....GluSerAlaGlyAlaGlyArgProThrProGlySerValArgAspArg 195
Qy 551 CGACACACAGAGAGAGTCTCTGGTGGAGAGAGCCAGAACCCAGAGATATATAGTCACTGC -- 495
Db 196 ArgArgGlyArgProGlyArgCysProGlyAlaLeuAlaGlyArgCysArgHisSerGly 215
Qy 494 TGAGGCTGGGCTGGCCACATGATCTCATCCAGAGAGACTCTCTATTGTCGATCTTAA 435
Db 216 AspGlyAspAlaArgAspValAlaGlyAspProValArgArgGlyTyrArgAlaAlaGly 235
Qy 434 AGACGGCCT.....GCTGTGAAGAGG 414
Db 236 ArgGlyProTyrArgSerPheGlyLeuProLeuPheMetGlnValAlaThrLeuProArg 255
Qy 413 AGCTGCTGTGGAGGCGAGTGTGCACACCGCACTGC.....TTCATGGAG 366
Db 256 AlaLeuAlaAlaValLeuLeuLeuValGluGlyLeuLeuHisArgLeuPheArgMetGln 275
Qy 365 CCACCTT...TGTCACAGGCTTCATTCTGATGAGGTCCAGCCAGCACTTTCGAGAA 309
Db 276 ProLeuGlnAlaPheGluGlnLeuAlaThrGlyLeuLeuGlyAsnAlaAlaGly 295
Qy 308 CAGACACAGTGCACCGAGTGTGATGCCATATACAGTCAAGTGTGTGGTTCGTCATTT 255
Db 296 AlaGlyThrAspLeuAspLeuGlyLeuAlaLeuAlaMetGlnAspLeuHisLeuHisGly 315
Qy 254 ....TTAGCTTCCGACACAGCAATATCA 231
Db 316 AlaGluLeuGlyGlnSerHisLeuGlnAlaAlaArgValAlaLeuGlyLeuAlaGlyAsp 315
Qy 230 ...GATCAATCATCCTCAGAGTCTTACACAG- 204
Db 336 LeuAspAlaProLeuHisGlyTyrGluGlnAlaArgGluLeuGlyGlnSerLeuPro 355
Qy 203 ACTCTGGAGTTTSCAGAGAGATATGTTGTTTATATATGATAATATATAGCCCTATGCG 144
Db 356 AlaProAlaAlaGlySerGlnLeuGlySerGlnLeuGlySerGlnLeuGlySerGlnLeu 366
Qy 143 AGTATCTCTCCGATTTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCTGAGTCTCT 106
Db 367 .....AlaValAlaGlyGlyProGlnGlyArgGlnGlyArgGlnGlyArgGlnGlyArg 382
Qy 107 TAGACTTTGGAGTAAGTGGCTGA...TCAGCCA 78
Db 382 alValGlyGlyGlnCysGlyTyrLeGlyPheLeuGlyLeuLeuLeuLeuLeuLeuLeuLeu 402
Qy 77 GCTTCACCCGAGTCCGAGTCTCAGAGAGATATCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 24
Db 402 lalPheAlaAlaArgCysArgGlnGlyArgGlnGlyArgGlnGlyArgGlnGlyArgGlnGly 422
Qy 23 TCTTTCTC.....TCTGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9
Db 422 rgAspLeuGlyCysGlyPheArgGlnGlyAlaGlyAlaGlyLeuTyrPro 438
```

RESULT 15

US-09-023 905A-2

Sequence 2, Application: US/39021935A

Patent No. 6473778

GENERAL INFORMATION:

APPLICANT: Roberts, Thomas M.

APPLICANT: King, Frederick J.

```
APPLICANT: Harris, David F.
APPLICANT: Hu, Eiding
APPLICANT: Spiegelman, Bruce
APPLICANT: Chan, Connie
TITLE OF INVENTION: Differentiation Enhancing Factors and Uses
FILE REFERENCE: EPN 021
CURRENT APPLICATION NUMBER: US/09/023, 905A
CURRENT FILING DATE: 1998-02-13
PRIOR APPLICATION NUMBER: US 60/038,191
PRIOR FILING DATE: 1997-02-14
NUMBER OF SEQ ID NOS: 39
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1129
TYPE: PRT
ORGANISM: Bos sp.
US-09-023-905A-2
Alignment Scores:
Pred. No.: 3,00121 Length: 1129
Score: 135.50 Matches: 109
Percent Similarity: 34.97% Conservative: 55
Best Local Similarity: 23.24% Mismatches: 163
Query Match: 5.89% Indels: 142
Dbs: 4 Gaps: 22
```

US-09-735-251-3 (1-1273) x US-09-023-905A-2 (1-1129)

```
Qy 62 CACTGGCGGTGAAGCTGGCTGACCGACCACTTACTCCAAAGTCTATTCTTCGGTTG--- 118
Db 609 HisLeuAlaValArgThrAlaAspGln-----ThrSerLeuHisLeuVal 623
Qy 119 -----CCAGACAGCAACTGGGAGATACTCG 145
Db 624 AspPheLeuAlaGlnAsnCysGlyAsnLeuAspLysGlnThrAlaLeuGlyAsnThrAla 643
Qy 146 CTAGCGGCGATAGTATTTCATTTCGACGACGCTTATTCTGCGAACTCCAGGAGTCT 205
Db 644 Leu--HisTyrCysSerMetTyrSerLysProGluCysLeuLysLeuLeuArgSer 662
Qy 206 GTTCAGAGAGCTGAGTGTGATTGAT-----CTGATCTACTGTGTCGGAAG 250
Db 663 LysProThrValAspValValAsnGlnAlaGlyGluThrAlaLeuAspLleAlaLysArg 682
Qy 251 CTAAAGATATACACGACACTTGACTTCTATGGCATTCACCTGGGTCCACT----- 301
Db 483 LeuLysAlaThrGlnCysGluAspLeuLeuSerGlnAlaLysSerGlyLysPheAsnPro 702
Qy 302 ---GTTCATCTT-----CTGCGAAAGTCTTCGCT 328
Db 703 HisValHisValGluTyrGluTyrAsnLeuArgGlnGluMetAspGluSerAspAsp 722
Qy 329 GAACCTGATAGAAACCGGACCTGTGCAGAAAGTGTGCTGCCATGAGAGAGTTCGCGGTG 388
Db 723 AspLeuAspLysProSerProIleLysLysGluArgSerProArgPro----- 739
Qy 389 TTGCACACTGCCCTGCACAGCAGCTCTCTTACAGGAGGCGGTCTTTAAGATGCTCAGC 448
Db 740 ---GlnSerPheCysHisSerSerSer-----IleSer 749
Qy 449 AATAGAGTGTCTGGATCAGATCATTTGTGCCACCCAGCCCTCAGCAGTACCCCTATT 508
Db 750 ProGlnAspLysLeu-----SerLeuProGlyPheSerThrPro----- 762
Qy 509 GCTCTTGGGCTTCTCCAGGACAGGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 568
Db 763 ----ArgAspLysGlnArgLeuSerTyrGlyAlaPheThrAsnGlnIle 777
Qy 569 ----GATAGCTTGGTGGCTGCTCACCCAGCCCTCGTCAATGCTGCTGCTGCTGCTG 616
Db 778 PheValSerThrSerThrAspSerProThrSerPro----- 789
```


GenCore version 5.1.6
Copyright (c) 1993 2003 CompuGen Ltd.

OM: protein - protein search, using sw mode:

Run on: November 12, 2003, 06:19:33 : Search time 20 seconds
(without alignments)
1371.458 Million cell updates/sec

Title: US-09-735-251-4

Perfect score: 2133

Sequence: 1 MS:SDWHLAVKLADQPLPK..... (SDWYGVWELMVRPSGV 416)

Scoring table: B:OSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 29308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 761*

1: Pir1*

2: Pir2*

3: Pir3*

4: Pir4*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	171	8.0	373	2 S54583	ubiquitin-like pro
2	164.5	7.7	536	2 B94547	probable ubiquitin
3	143.5	6.7	551	2 C94547	probable ubiquitin
4	141	6.6	528	2 T47147	saastic mucin like
5	141	6.6	671	2 T55568	hypothetical prote
6	139.5	6.6	1093	2 T35568	AP1 protein - hu
7	138.5	6.5	154	2 T34457	yeast Bsk2 homolog
8	138	6.5	872	2 S10201	hypothetical prote
9	138	6.5	992	2 A31668	hypothetical prote
10	136	6.4	1468	2 A44348	nucleoporin - rat
11	135.5	6.4	1162	2 JH0587	exo-alpha-sialidas
12	135.5	6.4	2715	2 T13049	eyelid - fruit fly
13	132.5	6.2	1952	2 T48814	hypothetical prote
14	130	6.1	1273	2 S58782	SEC3: protein - ye
15	130	6.1	1487	1 EDHEE1	immediate-early pr
16	127	6.0	398	2 S50507	excision repair pr
17	127	6.0	1487	1 EDHEE6	155k transcription
18	126.5	5.9	351	2 S50754	hypothetical prote
19	126	5.9	565	2 T39863	zinc finger protei
20	126	5.9	1147	2 T42627	ADP-ribosylation f
21	125.5	5.9	622	2 T15467	hypothetical prote
22	125.5	5.9	815	2 B56706	extracellular sign
23	125.5	5.9	990	2 T5668	nucleolar phospho
24	125	5.9	502	2 T18562	hypothetical prote
25	125	5.9	1613	2 S39057	protein BRG1 - hu
26	124.5	5.8	1676	2 A56558	anuclear primary
27	123.5	5.8	1744	2 A54973	tensin, cardiac mu
28	123.5	5.8	1792	2 A57075	tensin - chicken
29	123.5	5.8	3149	1 Q9BE5	BRP1 protein - hu

ALIGNMENTS

RESULT 1

S54583

ubiquitin-like protein DSK2 - Yeast (Saccharomyces cerevisiae)

N: Alternate names: protein YN8021.02; protein YMR276w

C: Species: Saccharomyces cerevisiae

C: Date: 08-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 24-Sep-1999

C: Accession: S54583; S59344

R: Pearson, D.; Bowman, S.

submitted to the EMBL Data Library, May 1995

A: Reference number: S54582

A: Accession: S54583

A: Molecule type: DNA

A: Residues: 1-373 <PEA>

A: Cross-references: EXBL:249704; NID:G8255340; PIDN:CAA89774.1; PID:G8255542; MIPS:YMR276w

A: Experimental source: strain AB972

R: Higgins, S.; Rose, W.D

submitted to the EMBL Data Library, April 1995

A: Description: The yeast ubiquitin-like proteins are involved in spindle pole body dur

A: Reference number: S59343

A: Accession: S59344

A: Molecule type: DNA

A: Residues: 1-108, R'110-295, R'297-373 <BIG>

A: Cross-references: EMBL:L40587; NID:G786149; PIDN:AAB07267.1; PID:G786151

C: Genetics:

A: Gene: SGD:DSK2

A: Cross-references: SGD:S0004889; MIPS:YMR276w

A: Map position: 13P

C: Superfamily: unassigned ubiquitin-related proteins; ubiquitin homology

F: 3-77: Domain: ubiquitin homology <UBH>

Query Match		8.0%	Score 171;	DB 2;	Length 373;
Best Local Similarity		22.4%	Pred. No. 0.0011;		
Matches	83;	Conservative	60;	Mismatches	127;
Indels	100;	Gaps	13;		
QY	63	LIYCGKLLKDDOTLDFYGIQPGSTVHVLRKSWPEPQKP	-----	101	
DB	44	LIYSGILKDDOTVFSYHIOGHSVHLV-KSQPKPTASAGANNATATGAAGTGATPN	-----	102	
QY	102	-----EPVDKVAAMREFVLHT-----	-----	132	
DB	103	MSSGSGAGFNPLADITSARYACYANPSADMFGPGGALNNDSNNQDELLRMENPIFQS	-----	162	
QY	133	----MLSNKESLJOITIVATPGLSS-DPIALGVLODKLFSVPADPNMLDTLVPAHALVN	-----	187	
DB	163	QWNEVLSNPQLDFWIOQNPOIQMGPOARQMLQSPMFRQMLTFPDM-----	-----	211	
QY	189	ALIVLVLHVSAGAPMPGCTDSSRSRSPSSYRMPGGLFEGLSDEDDFHNTSTPSS	-----	247	
DB	212	QMQFARMCDPNAGSGAGGASAPPA-----	-----	257	
QY	248	TPSSRPALSGYSGAGPRPIQSEALATALALASTPSSHTPTPTQCHSGSTSPMSGV	-----	307	

Db 258 TGNVAGTACGTAGNAGN-----TAANPFAS--LNPALNPPANAGNAASGCM 301
 QY 308 QSGTIP-ITNDLSQALQALQAGQSSQSQWQPCQLOQREKSLQDELSLPQCPPTGT 366
 DB 302 PAFDPALLASMFQPPVQ ASQAECTTPPEERVEHQRLQNCWFFFEFENVAALRRS-QG 359
 QY 367 SKCPWSSSLI 376
 DB 360 SVQGLDLSLL 369

RESULT 2
 B84549
 Probable ubiquitin-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: B84549
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 X.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: B84420; MUID:20083487; PMID:10617197
 A:Accession: B84549
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-536 <STO>
 A:Cross-references: GB:AE002093; NID:94584343; PIDN:AAD25138.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g17190
 A:Map position: 2

Query Match 7.7% Score 164.5; DB 2; Length 536;
 Best Local Similarity 24.0%; Pref No. 0.0046;
 Matches 99; Conservative 49; Mismatches 140; Indels 125; Gaps 19;

QY 39 SISFLKLIAGKIQESVDPDELILYCGRKLKDDQTLDFYIQPGSTVHLVKSPPEPD 98
 DB 39 TVESFKELIA--QNSDVPANQORLLYKGRILKDDQTLTLLSYGLQADHTIHMVRGSAFSSA 95
 QY 99 QKPEPV----- 104
 DB 96 PPPAPASQTTAPSVTRGVGSDNSSNLGGASPGESLPGLGFNPLGGGNAMSLFGAGLP 155
 QY 105 DKVAAMREFRVLTALHSSSSSYREAVFKLSNKESLDQIIIVATPGLSSDDPIALGVLDQK- 163
 DB 156 DLVQFQQLAQNPNYIFDMYN-TPAIQNANPNPEFMRSMIMNPNQMR-----LVDRN 207
 QY 164 -DLFSVPADPNML-DTLVPA-HPALVNAIVLVHSVAGSAPMPGCTDSSSSSPSSSSYRDM 220
 DB 208 PELGHVINDPSILRQTLLEAARNPELAKEM-----MRNTD---RAM--SNIESM 250
 QY 221 PGGF-----PEGLSDDEDQFHNTSTPSS:TPSSRPASLGYSGAGPRPITQSELATA 275
 DB 251 PEGNVLRRYENVQEE-----PLMNATMSGNAGNTGNPNPFAALLGNOCVT-TQGSDA 303
 QY 276 LALASTPESSESH-----PTPGTQGHSSG--TSPMSSGV----- 307
 DB 304 SNNSTPNAGTGTIPNANPLNPMGATGGOTTAPRTNVGDDARSPLGLGLGLGLGLGL 363
 QY 308 -----OSGTPITNDLFSQALQALQAGQSSQSQWQ-----POLQQLRD 347
 DB 364 GGLGLGADSPGATGPASQL-SQLLQNPATISQMQSVFSNPQYMNQMLSNPOLRSMLD 422
 QY 348 MGIDDELRLRP 359
 DB 423 SNPGLPEMNQNP 434

RESULT 4
 I47141
 Gastric mucin (clone PGM-2A) - pig (fragment)
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 03-Nov-2000
 C:Accession: I47141; S55315
 R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
 Gastroenterology 106, 200, 1994
 A:Title: Pig gastric mucin: isolation and characterization of a cDNA clone with a nov
 A:Reference number: I47141; MUID:94102478; PMID:7506218
 A:Accession: I47141
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-528 <TUR>
 A:Cross-references: EMBL:U10281; NID:g915205; PIDN:AAC48526.1; PID:g915208
 R:Turner, B.S.; Bhaskar, K.R.; Hadzopoulou-Cladaras, M.; Specian, R.D.; LaMont, J.T.
 Biochem. J. 308, 89-96, 1995
 A:Title: Isolation and characterization of cDNA clones encoding pig gastric mucin.
 A:Reference number: S55315; MUID:95275264; PMID:7755593
 A:Accession: S55315
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-528 <TU2>

RESULT 3

B84549
 Probable ubiquitin-like protein [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: B84549
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 X.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

A:Cross-references: GB:J10281; NID:9915205; PIDN:AA048526.1; PID:9915209
 C:Superfamily: pig submaxillary mucin

Query Match 6.6%; Score 141; DB 2; Length 528.
 Best Local Similarity 22.3%; Pred. No. 0.14;
 Matches 77; Conservative 51; Mismatches 159; Indels 58; Gaps 11;

QY 118 TALHSSSSVREAVFKMLSKESLQGLIVATPGLSERIALGVLCXKELFVFAADPNYLDI 177

DB 178 TVSQPSSSSSPPTSSVGVCTSSSAVPTTSITVQPSSSSSVPTTSATSVKSSSSST 236

QY 178 LVP-----AHPLAVALVILVSVAGSAVPTT-----DSKSSKSSSSSYADY-PGG 223

DB 237 PISTTSVPSSSSSAFTTSA-SVQSSSSSTSTSTTSVPSSSSSAFTTSA-SVQSS 296

QY 224 FLEGLSDDEDDPHNTRTPSSSTSSHPASL---YSGAARPPPT-SBLA-LALAN 280

DB 297 SSSPPISSITSVQPSSSSSSPPTSTTSVQPSSSSGAPTTSATSVQPSSSSSSP 356

QY 281 TPSSSHTPTPGT---GCHSSGTSFVSSGV-----CSGTF--T 314

DB 357 QPSSSSSPTTSTTSVQPSSSSAFTTSA-SVQPSSSSSVPTTSATSVKSSSSSTPIPT 416

QY 315 NDLFSQALCHALQASQPSLQSQWQPCQLQKRWYIQDELRLRQCR---PPVGTISKQ 369

DB 417 TTSVQPSSSSSVPTTSATSVQSSSSSTSTSTTSVQPSSSSSATTSATSVQPSSSSSP 426

QY 370 PWSSSLLEPHELPASP-----EPFASGRYVHWELMKVDFSS 407

DB 477 PISTTSVPSSSSSPPTSTTSVQPSSSSAFTTSA-SVQPSSSSAFTTSA-SVQPS 519

RESULT 5

102504
 Hypothetical protein A2938410 (imported). Arabidopsis thaliana

N:Alternate names: hypothetical protein T19C21.10

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 05-Mar-1999 #sequence_revision 15-Mar-1999 #text_change 16-Feb-2001

C:Accession: T02504; F84804

R:Rounsley, S.D.; Lin, X.; Kechum, K.A.; Graczyk, M.L.; Branden, K.C.; Sykes, S.N.; Kaul

submitted to the EMBL Data Library, August 1999

A:Description: Arabidopsis thaliana chromosome 1: BAC T19C21 genomic sequence.

A:Reference number: 214676

A:Accession: T02504

A:Status: translated from GB/EMBL/DBE

A:Molecule type: DNA

A:Residues: 1-671 <RCU>

A:Cross-references: EMBL:AC004684; NID:9915205; PID:9915209

A:Experimental source: Cultivar 'Colombino'

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Lin, X.; Graczyk, M.L.; Branden, K.C.; Sykes, S.N.; Kaul

M.J.; Koo, H.; Moffat, K.S.; Graczyk, M.L.; Branden, K.C.; Sykes, S.N.; Kaul

euss, D.; Nierman, M.C.; White, D.; Graczyk, M.L.; Branden, K.C.; Sykes, S.N.; Kaul

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 1: BAC T19C21 from Arabidopsis thaliana.

A:Accession number: A84420; MUID:20080429; PMID:10511197

A:Accession: F84804

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-671 <STO>

A:Cross-references: GB:ACC02093; NID:9915205; PIDN:AA048526.1; GB:GB:GX00139

C:Genetics

Query Match

Best Local Similarity 6.6%; Score 141; DB 2; Length 671.

Matches 81; Conservative 40; Mismatches 130; Indels 112; Gaps 14;

QY 99 QKPEPVKVAAMREFRVLTALHSSSSSYRAVFKMLSKESLQGLIVATPGLSERIALGVLCXKELFVFAADPNYLDI 158

DB 253 QAVDFSSREAVKDEIV-----TLVERTSNQLKLMQVITS----- 289

QY 159 VLQDKDLFSVFAOPN-MLDTLVPAPALVNAIVLVH----- 194

DB 290 --GDDELLORGJDLNDSQIILLAKHDAIASGSLPVQASGSLSVQASKPADSSPKSSEA 347

QY 195 ----SVAG-SAPVPGTGDSSSRMPSSSYRMPGGFLFEGLSDDEDDPHNTR--STPSSS 247

DB 348 KSSSS-AGSSSPITATVSTGKSPIDEV-----EEDEFAQLARRHSKPPAS 395

QY 248 TSSRRPASVGYSAAG-----FRPITOSE-----LATALALASTPSSSHTPT 290

DB 396 V-TTDTTS-ESFNAAASNALALALPDPFPVNTTKQDMIDLLSITLCTPSTPPAPSSQPS 454

QY 291 PGTQGH-----SGTSPMSSGVCSGTPTTNDLFSQALQALQ---ASQOP 332

DB 455 PPPPASGDNTH:YPOQPFRFSYVAPWACQCPQPOAQQGTSQHQQOQGYSPQPH 514

QY 333 SLGSRQPCQLQRLDMG:QDDELSLRPCRPVGTSTKQPMSSSLSEPHLPASPEPPASC 392

DB 515 SQOQGYSLQCC-----PQOQGYSGQPOAQVQMP--STRPONVEYPPPPMASTSA 566

QY 393 RGY 395

DB 567 NAY 569

RESULT 6

138533

AF17 protein - human

C:Species: Homo sapiens (man)

C:Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 01-Dec-2000

C:Accession: 138533; T43469

R:Prasad, R.; Leshkowitz, D.; Gu, Y.; Alder, H.; Nakamura, T.; Saito, H.; Huebner, K.;

Proc. Natl. Acad. Sci. U.S.A. 91, 8107-8111, 1994

A:Title: Leucine zipper dimerization motif encoded by the AF-17 gene fused to ALL-1 (ML

A:Reference number: 138533; MUID:94336895; PMID:8058765

A:Accession: 138533

A:Molecule type: mRNA

A:Residues: 1-1093 <RES>

A:Cross-references: EMBL:U07932; NID:G532761; PIDN:AAA21145.1; PID:G532762

R:Blöcker, H.; Boecker, V.; Brandt, P.; Vewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 22551

A:Accession: T43468

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 670-1093 <AAA>

A:Cross-references: EMBL:U07932

A:Experimental source: adult testis; Clone DKFZp434K0227

C:Comment: This gene is known as an occasional fusion partner in chimeric proteins cre

C:Genetics

A:Gene: GB:MUT6; AF1

A:Cross-references: GDB:385409

A:Map position: 17q21.17q21

A:Note: DKFZp434K0227;

Query Match

Best Local Similarity 6.5%; Score 139.5; DB 2; Length 1093;

Matches 108; Conservative 48; Mismatches 181; Indels 167; Gaps 20;

QY 11 KLADQP-T--PKSLRLPETELGEYSJGGYSISFLKQLIAG-----KLOESVP 56

DB 590 RLRSRPTSTSLPSSSSASISTQV--FSLAGSTFS:PSHTFOTPMGAVNPLLSQAESHT 647

QY 57 DPESLDLIVGGRKDDQDLDFYGVQPGSTVHV-----LRKSWPEPDQKPE 102

DB 648 EPDLEDCSFCRCSTSPQESLS--SMSPISLPAFDQTASAPCGGQLDPAAPGTTNMEQ 705

QY 103 PVDK-----VAMREPRVLTALHSSSSSYREAFKMLSKESLQGLIVATPGLSSTPI 155

DB 706 LJEKQGTGEAGVNVEMKALHALQENORLQEQILSLTAKKERLQILNV----- 755

QY 156 ALGVLCJCKLFSYFAFDNMLDT--VPAFPALVNAIVLVHSVAGSAFMPQTO---SSSRSM 212

A:Note: sequence extracted from NCBI backbone (NCBI:79200, NCBI:79212);
C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3C

Query Match: 6.5%; Score 139; DB 2; Length 992;
Best Local Similarity 22.5%; Pred. No. 0.5;
Matches 71; Conservative 35; Mismatches 116; Indels 94; Gaps 15;

QY 177 TLVPAHPALVNAIVLHLS --VAGSAPYCT 255
DB 461 TVERAHPVEMPMVLHCHPPVPPVPPKPTPTTSRRPGACVYDDEVLVVDVETTE 520
QY 206 DSSSSMSSSSYRDVPGFLEGLSD... DEDEHENTRTNTSSATPPSPASLG 257
DB 521 DSSSVCPKPHKQCDGQFQSGREQKAAATVSPSTGDFAPVPAAGPAPAPAA 580
QY 258 YSG----AGRPITQSELATALASTPSSSHPTPGTQ..... 294
DB 581 PPAAGPAPAGP-PAAGRPILAPLS-AGPAAAGPILVTPPSARPVAPVVRMPVREQL 638
QY 295 GHSSGTSMS--GVCSGTPTINDLFSQALCHALQASQFSL-QSCWQPLQQJ--RQNG- 349
DB 639 POSTGRKQCENEMAGRE:TO--MQQPSHLSGATQPTTPRSWAPSVYCALSUMDAGK 695
QY 350 ----ICDDELSPPCPPVGTGKQSSSLMEPEH... --EUPASPPFASCPY 395
DB 697 AQPIESSHLSMSPTQPSHEQP... EYEDCAFLDLSLHPTVAAGPAPAPVQGYQE 752
QY 396 ----CPKEALMKVPP 406
DB 753 PPAPCPVQGYQEP 466

RESULT 10
A44345
nucleoporin . rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Jun-1993 #sequence_revision 14 Nov 1994 #text_change 63 Mar-1995
C:Accession: A44345
C:Superfamily: R:Sukenaga, J.; Blobel, G.
C: 72, 29-18, 1993
A:Title: A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and
A:Reference number: A44345; XMOD:19313735; PMID:8422679
A:Accession: A44345
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-1468 <SW>
A:Experimental source: liver
A:Note: sequence extracted from NCBI backbone A:BIN:12145, NCBI:P:123338;

Query Match: 6.4%; Score 139; DB 2; Length 1468;
Best Local Similarity 23.4%; Pred. No. 1;
Matches 90; Conservative 45; Mismatches 145; Indels 139; Gaps 17;

QY 31 GYSILG---GYSISLKQLIAGKQESVPGTELLGAGVQKLI --KCO 73
DB 1073 GGFTEGVDSALSPSMFVLQTEKQCEPVTSLVFEKKAQNEBPXKQVPSFENSE 1132
QY 74 QTLDFYGIQGSTVHLRKNPEPCKPE..... --PVQVAMREFRVHTALHS 122
DB 1133 QTKDESSSKPTFSFVAKPSVKESDCLAKAFAPNQNTTTCQAAKPAFSLNSSSS 1192
QY 123 SS-----SYREAVKMLNKESLDQIIIVATPLSSSPIALGVQCKLFSVPADPMJLIL 178
DB 1193 SSTPATSSASIFGS-STSSSSFPVAAFFVGGASNPVS..... --SSAFNSAESSTIS 1242
QY 175 VP-----AHPALVNAIVLVHVSAGAP --VGTSSSSSYRPSVYDPMQGFEEGLS 230
DB 1243 QPLIFFQDCKPATTS..... --STASAPFVFGVQASSNSTVSS --GTFGATT 1298
QY 231 DEDDQHPNTRTPSS----- --STFSPPA SLYSVAAGPPPTQSELA 274
DB 1289 -----TSSSSGSFFVFGTGHGAPASPAFGANQFTTFQGSAGAQFPSPFGHSS 1339

QY 275 ALALAS-----TPSSSHPTPT-----PGTQGHSSGTSMSVQSGTPTINDLP 318
DB 1340 STALFSAGSQVPPPTPTGTVSSSQPPVFGQPSQSAFGSGTANASSVFGSGSTTNFNF 1399
QY 319 SQ-----A-QHALQASQPS 333
DB 1400 TNNPSPGVFTGASPSPTPAAAAQPS 1424

RESULT 11

JH0557
exo-alpha-stalidase (EC 3.2.1.18) - Trypanosoma cruzi
N:Alternate names: neuraminidase
C:Species: Trypanosoma cruzi
C:Date: 10-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 11-Jan-2000
C:Accession: JH0557
R:Peirard, M.E.A.; Meiss, J.S.; Ortega-Barria, E.; Matzilevich, D.; Prioli, R.P.
J. Exp. Med. 174, 179-191, 1991
A:Title: The Trypanosoma cruzi neuraminidase contains sequences similar to bacterial ne
A:Reference number: JH0557; MUID:91277609; PMID:1711561
A:Accession: JH0557
A:Molecule type: DNA
A:Residues: 1-1162 <PE>

A:Cross-references: GB:XA1732; NID:9162302; PID:9162303
A:Note: the authors translated the codon TCT for residue 45 as Cys
C:Comment: This protein plays a role in parasite-host cell interaction.
C:Superfamily: trypanosomatid-specific surface antigen
C:Keywords: glycoprotein; glycosidase; hydrolase
F:394/Binding site: carbohydrate (Asn) (covalent); #status predicted

Query Match: 6.4%; Score 135.5; DB 2; Length 1162;
Best Local Similarity 21.8%; Pred. No. 0.9;
Matches 98; Conservative 51; Mismatches 189; Indels 111; Gaps 18;

QY 6 WELAVKLACQLTPKSLRLPETELGEYSJGYSISFLKLIACKLQESVDPDELID--- 62
DB 499 YELVLTMAN-----XIGSVYIDG-----ELLEGSGTVPDGTPDISH 537
QY 63 -----LYCGRKLDKDDTLDFYGIQV--GSTVHVLRKSWPEPD 98
DB 538 FVVGKYSQDVFYISHV--VNNVLLNRPQLNTEIRTLFLSODLIGTEAHWDSSDSSAH 597
QY 99 QKPE-PVQKVAAMREFRVHLTHLSYSSRYEAVPKMLNKESLDQIIIVATPLGSSDPIAL 157
DB 598 STPTPADSSSAHSTPTPDSSAHSSTPTPADSSAHSSTPTPDSSAHSSTPTPADSSAH 657
QY 158 GYLQDKQLFSVPADPNM-LDTLVPAHPALVNAIVLVHVSAGSAPMPCGTSSSRMPS-- 214
DB 658 GTPSTPVSAAHSTPTPTVDSSAHSPTS-----TPVDSSAHSSTPTPADSSAHSSTPT 711
QY 215 --SSYRDPVSGFLEGLSCDEDDFHFNTRSTPSSRPSALGYSAGAPRITQSEL 272
DB 712 VDSSAHSSTPTPADSSAHSSTPTPADSSAHSSTPTPADSSAHSSTPTPADSSAHSSTPT 761
QY 273 ATALALASTPSSSHPTPTGT-----QGHSSGTSMSVQSG--GTPITNDLFSQALQHAL 326
DB 762 SSAHSTPTPADSSAHSSTPTPADSSAHSSTPTPADSSAHSSTPTPADSSAHSSTPTPDVS 810
QY 327 QASGQPS-QSQWQPLQQLQRMGIQDDEL-SURPCRP-----PVGTSKOPWSSSLLEPHEL 382
DB 811 SAHGTPTPT-----VDSSAHSSTPTPDSSAHSSTPTPDSSAHSSTPTPDVS-----AHST 851
QY 383 PASP-EPASCSGYPWEALMKVPSLPV 410
DB 852 PSTPADSSAHSSTPTPADSSAHSSTPTPV 880

RESULT 12

T13049

eyelid - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: T13049

R.Treisman, J.E.; Luk, A.; Rubin, G.M.; Heberlein, U.
submitted to the EMBL Data Library, March 1998
A:Reference number: Z17592
A:Accession: T13049
A:Status: preliminary; translated from OE/EMBL/CDEN
A:Molecule type: mRNA
A:Residues: 1-2715 <TRE>
A:Cross-references: ENH1.APOS3091; NID:3258122; P103298122; P10N.AAC06254.1
A:Genetics:
A:Gene: old
A:Cross-references: FlyBase:FBgn:0003013
C:function:
A:Description: could act as a transcription factor antagonistic to the WG pathway
C:Keywords: DNA binding

[illegible]

1129 -----TTNSTGSSNSGSDPAPFKAAPNAALLTPYQEGSGPTGASGSP --DYAT 1178
 QY 275 ALALASTP-ESSSHPTPTSC -----CHSSTSP ----- 302
 DE 1179 AGQKRIPSQNPOTPHPGAAAVAAGNTIVSNPFEDTAAAGTGGGTGPGQGGPG 1218
 QY 303 -----MSSGVQSGTPTINLFSALLNALCASG -----RELSSQWPGCQ 344
 DE 1239 GAASGGAGAVAGVGGCGPHPPFPSPHTAAQCAAGHCCQPFQHQRPULPFPFPQQQ 1298
 QY 345 LRDMLQDDELSLRPPAPVGTGKQWSSLLRPHIELPSTP -----PASTRGYCP 397
 DE 1299 ---GCGCGG-----PFAVQVQETAPQLRQGVETSLIVRFAAGATPEP 141

RESULT 13
 T48814
 hypothetical protein I56.220 [transport] Neurospora crassa
 C.Species: Neurospora crassa
 C.Date: 05-May-2003 sequence_revision 55-May-2003 text_change 55-May-2003
 C.Accession: T48814
 R.KSchulte, C.J. Aigui, V.J. Hoheisel, C.J. Blardt, P.J. Farnham, B.J. Holland, R.J. Nyakatura,
 submitted to the Protein Sequence Database, April 2003
 A.Reference number: Z24541
 A.Accession: T48814
 A.Status: preliminary
 A.Molecule type: DNA
 A.Residues: 1-1952 <SCH>
 A.Cross-references: EMBL:AL35322; GSPDB:GN03112; NCSP:I56.220
 A.Experimental source: cosmid contig I56, strain 74
 C.Genetics:
 A.Gene: NCSP:I56.220
 A.Map position: 2
 A.Introns: 28/3

Query Natch 6.2%; Score 132.5; DE 2; Length 1952;
Best Local Similarity 22.4%; Pred. No. 2.8;
Matches 79; Conservative 45; Mismatches 140; Indels 89; Gaps 16;
CY 81 ICGSTVHLKRSNFEQDKP- EFTQKVA-MREFRVHIALISSKRYEAEVFNLSN 136

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 06:28:53 : Search time in seconds
without alignment:
1334.173 Million cell updates/sec

Title: US-09 735 251 4
Perfect score: 2133
Sequence: 1 NLSLWHLAVKALQCFUPK SRYGCPWELKVFSLPVL 41

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 203000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt_41*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	170	8.0	373	1	DSK2_YEAST
2	143.5	6.7	555	1	DAB2_HUMAN
3	141.5	6.6	555	1	DAB2_MOUSE
4	139.5	6.5	1093	1	ATF7_HUMAN
5	138.5	6.5	354	1	YAC9_SCHPO
6	138	6.5	992	1	ENH4_EBV
7	136	6.4	468	1	N553_RAT
8	135.5	6.4	1662	1	TYNA_TPOVR
9	133.5	6.3	907	1	ATF7_HUMAN
10	130	6.1	1273	1	W511_YEAST
11	130	6.1	1457	1	1004_HUMAN
12	129	6.0	598	1	N541_CACAE
13	127	6.0	398	1	R242_YEAST
14	127	6.0	1487	1	1014_HSVB
15	126	5.9	565	1	SRL1_SCHPO
16	125.5	5.9	815	1	MC27_HUMAN
17	125.5	5.9	324	1	1K52_HUMAN
18	125	5.9	676	1	MUC1_MESAL
19	125	5.9	707	1	BC16_MOUSE
20	125	5.9	5120	1	PCO2_HUMAN
21	124.5	5.8	1676	1	APSA_EMENT
22	124	5.8	1233	1	HKH4_HUMAN
23	123.5	5.8	1744	1	TENS_CHICK
24	123.5	5.8	3149	1	TECU_EBV
25	123	5.8	1844	1	PCOL_TYVYC
26	122.5	5.7	1210	1	AP4_HUMAN
27	122	5.7	1305	1	GAK_RAT
28	122	5.7	1487	1	MS3_YEAST
29	122	5.7	1844	1	PCOL_TYVY
30	121.5	5.7	706	1	BC16_HUMAN
31	121.5	5.7	3421	1	TECU_HSVB
32	121	5.7	2161	1	SHK1_HUMAN
33	121	5.7	3530	1	MY15_HUMAN

34	120.5	5.6	1465	1	NCO2_RAT
35	120.5	5.6	5147	1	PCO2_HUMAN
36	120	5.6	770	1	DAB2_MOUSE
37	120	5.6	1983	1	TF20_MOUSE
38	119	5.6	1367	1	AMH_YEAST
39	119	5.6	1794	1	YDC9_SCHPO
40	118.5	5.6	1206	1	P3PA_HUMAN
41	118.5	5.6	1112	1	GAK_HUMAN
42	118	5.5	269	1	NO20_MESAL
43	118	5.5	145	1	TOB1_HUMAN
44	118	5.5	971	1	RJHD_HUMAN
45	118	5.5	2404	1	SON_MOUSE

ALIGNMENTS

RESULT 1	DSK2_YEAST	STANDARD	PRT: 373 AA
ID	DSK2_YEAST		
AC	P48510		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	01-FEB-1996 (Rel. 33, Last sequence update)		
DT	15-SEP-2003 (Rel. 42, Last annotation update)		
DE	Ubiquitin-like protein DSK2		
GN	DSK2 OR SHE4 OR YMR276W OR YMR021.02		
OS	Saccharomyces cerevisiae (Baker's Yeast)		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
CC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.		
CX	NCBI_TaxID:4932;		
RN	[1]		
R2	SEQUENCE FROM N.A.		
RC	STRAIN=S288C		
RX	MEDLINE=96281973; PubMed=9628268;		
RA	Biggins S., Ivanovska I., Rose M.D.;		
R7	"Yeast ubiquitin like genes are involved in duplication of the		
RT	"Microtubule organizing center."		
RL	J. Cell Biol. 133:1331-1346(1996).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=S288C / AS472;		
RX	PubMed=9169872;		
RA	Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,		
RA	Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,		
RA	Jagels K., Lye S., Moule S., Odell C., Pearson D., Rajandream M.A.,		
RA	Rice P., Skelton C., Walsh S., Whitehead S., Bartell B.G.;		
RT	"The nucleotide sequence of Saccharomyces cerevisiae chromosome		
RT	XIII."		
RL	Nature 387:90-93(1991).		
CC	FUNCTION: INVOLVED, WITH RAD23 IN SPINDLE POLE BODY DUPLICATION.		
CC	SUBCELLULAR LOCATION: Nuclear (Probable).		
CC	SIMILARITY: TO S. POMBE SPAC26A3.16.		
CC	SIMILARITY: Contains 1 ubiquitin-like domain.		
CC	SIMILARITY: Contains 1 UBA domain.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/		
CC	or send an email to license@sib-sib.ch).		
CC	EMBL; L40587; AAC37267.1; ..		
DR	EMBL; Z49704; CAAB9774.1; ..		
DR	PIR; S54583; S54583.		
DR	HSSP; Q15843; INDD.		
DR	SGD; S0004889; DSK2.		
DR	InterPro; IPR006636; Still.		
DR	InterPro; IPR000449; UBA_domain.		
DR	InterPro; IPR000626; Ubiquitin.		
DR	Pfam; PF00627; UBA; ..		
DR	Pfam; PF00240; ubiquitin; 1.		

```

DR PRINTS: PR00348; UBIQUITIN.
DP SMART: SM00727; ST11.2.
DR SMART: SM00165; UBA.1.
DR SMART: SM00233; URO.1.
DR PROSITE: PS00030; UBA.1.
DR PROSITE: PS00299; UBIQUITIN.1.
DR PROSITE: PS00053; UBIQUITIN.2.
KW Nuclear protein.
FT DOMAIN 1 76 UBIQUITIN LIKE
FT DOMAIN 327 371 USA
FT CONFLICT 109 109 R -> A (IN REF. 2).
FT CONFLICT 296 296 S -> A (IN REF. 2).
SQ SEQUENCE 373 AA: 25EPFR29925K UFE CP-054.

Query Match      8.04; Score 170; E: 1e-17; Length 373;
Best Local Similarity 22.74; Pred. NO. 0.13;
Matches 83; Conservative 60; Mismatches 123; Indels 100; Gaps 13;

QY 63 LIYCKRLKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 106
DB 63 LIYCKRLKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 106
QY 67 GRKLRKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 116
DB 67 GRKLRKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 116
QY 110 HTALHNSSTREAV - FKMUNKESLCQI:VATPGLSDFIALGVLCOKLFSVPAD 171
DB 107 HHAVEEETIAKDTTHPAFYGVCGEGNHRFVA:KTAQAEPV---ILCLRDLFQIYE 163
QY 172 PNMCTIL-----VPAHPALVNAIV-----LVLSVAGSAPM--PGTSSSRS 211
DB 164 LKQREELKKAQKDKCECAVYOTILEEDVEPVYCYIVFE--AGHEPIRDPETEENIYQ 221
QY 212 KPSSSYVDMPOGFLFEGLSDEDDPHNTRSPSSSTPSRPSASUGYGAAGPRPTOSE 271
DB 222 VPTSQKK-----EGVYD-----VPKSQPV-----AVTGLE 247
QY 272 LATALALASTPSSSHPTPTGTCQ-----HSSGTSFMSGVSGTPTINDLFSQALQAL 326
DB 248 LFGCM- -STPPTDITSPPTATGDAPIFSSQTLPASADVFSSVP-----LGTAAVPGY 300
QY 327 QASGQ PSLOSQWQPOLQJRDIMGIDDELRLCRPPVGTSGKOPWSSSLEPHELPA 384
DB 301 VAMGAVLPSEWGO-QPLVQQQWYMGARPPVAVQVMPGCAQPIANG-QPGLFPATQOPWPTVA 358
QY 385 SPEIPASCRGVCPEALMKVPPSL 408
DB 359 GQFPAA---FMPTCTVMP:PAAM 379

Query Match      6.74; Score 143.5; DB 1; Length 555;
Best Local Similarity 22.74; Pred. NO. 0.13;
Matches 87; Conservative 51; Mismatches 143; Indels 103; Gaps 19;

QY 67 GRKLRKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 116
DB 67 GRKLRKDDQTLDFYGIQPGSTVHLVLRKSWPEPOCKPEPV-----DKVAAMREFRVL 116
QY 110 HTALHNSSTREAV - FKMUNKESLCQI:VATPGLSDFIALGVLCOKLFSVPAD 171
DB 107 HHAVEEETIAKDTTHPAFYGVCGEGNHRFVA:KTAQAEPV---ILCLRDLFQIYE 163
QY 172 PNMCTIL-----VPAHPALVNAIV-----LVLSVAGSAPM--PGTSSSRS 211
DB 164 LKQREELKKAQKDKCECAVYOTILEEDVEPVYCYIVFE--AGHEPIRDPETEENIYQ 221
QY 212 KPSSSYVDMPOGFLFEGLSDEDDPHNTRSPSSSTPSRPSASUGYGAAGPRPTOSE 271
DB 222 VPTSQKK-----EGVYD-----VPKSQPV-----AVTGLE 247
QY 272 LATALALASTPSSSHPTPTGTCQ-----HSSGTSFMSGVSGTPTINDLFSQALQAL 326
DB 248 LFGCM- -STPPTDITSPPTATGDAPIFSSQTLPASADVFSSVP-----LGTAAVPGY 300
QY 327 QASGQ PSLOSQWQPOLQJRDIMGIDDELRLCRPPVGTSGKOPWSSSLEPHELPA 384
DB 301 VAMGAVLPSEWGO-QPLVQQQWYMGARPPVAVQVMPGCAQPIANG-QPGLFPATQOPWPTVA 358
QY 385 SPEIPASCRGVCPEALMKVPPSL 408
DB 359 GQFPAA---FMPTCTVMP:PAAM 379

RESULT 2
DAB: HUMAN
AC 075533; Q9NYA8;
DT 28-FEB-2003 (Rel. 41, Created:
DT 28-FEB-2003 (Rel. 41, Last sequence update:
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DB Disabled homolog 1.
QY DAB:
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catartida; Hominidae; Homo
CX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99009143; PubMed=979077;
RA Lambert de Rouvroit C., Goffinet A.M.;
RT Cloning of human DAB1 and mapping to chromosome 5 (11-p11.1);
PL Genomics 53:246-247(1998);
SN 12;
RP SEQUENCE FROM N.A.
RA Fazili Z., Sur W., Xu X.;
RT "Aberrant disabled-1 expression in tumors";
SL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases
-- FUNCTION: Adapter molecule functioning in protein development. By
-- similarity

```

```

RESULT 3
DAB1_MACFA STANDARD; PRT: 555 AA.
AC Q9BGX5;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Disabled homolog 1.
GN DAB1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
CX NCBI_TaxID=9541;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=Frontal cortex;
RA Csada N., Hida M., Kusuda J., Tanura K., Isaki K., Hirai Y., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.
RT "Isolation of full length cDNA clones from macaque brain cDNA
RT libraries."
RL Submitted (FEB-2001) to the EMBL/GenBank/CCDC databases.
CC -!- FUNCTION: Adapter molecule functioning in neural development. By
CC similarity).
CC -!- SUBUNIT: Associates with the SH2 domains of Src, Fyn and Abl (by
CC similarity).
CC -!- DOMAIN: THE PID DOMAIN SPECIFICALLY BINDS TO THE ASN-PRO-XAA-
CC TYR(P) MOTIF FOUND IN MANY TYROSINE-PHOSPHORYLATED PROTEINS.
CC -!- PTM: Phosphorylated on Tyr-198 and Tyr-220 upon insulin induction
CC in embryonic neurons (by similarity). Also phosphorylated on Ser-
CC 491 independently of insulin signaling (by similarity).
CC -!- SIMILARITY: Contains 1 PID domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
EMBL; AB055282; BAB21906.1;
DR InterPro; IPR006C23; PTB_PID.
DR Pfam; PF00640; PID.1.
DR SMART; SM00462; PTB.1.
DR PROSITE; PS01179; PID.1.
KW Developmental protein; Neurogenesis; Phosphorylation
FT DOMAIN 36 189 PID.
FT MCC_RES 198 198 PROSER-PYLATED BY SIMILARITY
FT MOD_RES 220 220 PROSER-PYLATED BY SIMILARITY
FT MOD_RES 491 491 PHOSPHORYLATION BY SER 491
FT SIMILARITY
SQ SEQUENCE 555 AA; 59920 MW; A3B802A1D3C4E5E1 CRD54.
Query Match 6.6%; Score 141.50; E-Value 1.0e-55;
Best local Similarity 21.8%; Pred. No. 517;
Matches 83; Conservative 52; Mismatches 150; Indels 95; Gaps 18;
QY 67 GRKLDDQTLDFYGIQPGSTVHVLRKSWDEPKKPEPV...DKVAMSEFRVL 116
Db 57 GDLKCDMMKLGKVAGA...RKSEHKQKLELTISFGKIFDEKCALQHHA 110
QY 117 HTALH-SSSSYRAVPMK-SNKSLEDCQIVATPG-SSODIALGV-CKDLGVFAFDPNKL 175
Db 111 HEISY-ASYOTTHRAFYGACGKGNHFRVAIKTAAGSEV...LCLRLFLVLELQK 167
QY 176 DTL-----VPAHPLVNAIV-----VLHVSAGAPK-PCTSSSPSPSS 215
Db 169 EELEKAKQDKQCEQAVYQTI-EEEDVEDPVYQVYFE--AGHEFIRDPETRENIYQVPT 225
QY 216 SYRDMPGFLFEGLSDEDDDFHENTSTPSSSTPSRRPASLGYSGAGPRPTQRELA 275
Db 216 SYRDMPGFLFEGLSDEDDDFHENTSTPSSSTPSRRPASLGYSGAGPRPTQRELA 275

RESULT 4
AF17_HUMAN STANDARD; PRT: 1093 AA.
ID AF17_HUMAN
AC P55198;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE AF-17 protein.
GN MLLT6 OR AF17.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
CX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=9433695; PubMed=8058765;
RA Prasad R., Leshkowitz D., Gu Y., Alder H., Nakamura T., Saito H.,
RA Huebner K., Berger R., Croce C.M., Canaan E.
RT "Leucine zipper dimerization motif encoded by the AF17 gene fused to
RT A2-1 (MLL) in acute leukemia."
RL Proc Natl Acad Sci U.S.A. 91:8107-8111 (1994).
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DISEASE: INVOLVED IN ACUTE LEUKEMIAS BY A CHROMOSOMAL
CC TRANSLOCATION T(11;17)(Q23;Q21) THAT INVOLVES MLLT6 AND MLL/HRX.
CC THE RESULT IS A ROGUE ACTIVATOR PROTEIN.
CC -!- SIMILARITY: Contains 2 PHD-type zinc fingers.
CC -!- SIMILARITY: HIGH TO AF10.
CC -!- DATABASE: NAME=Acias Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/Chromocancer/Genes/AF17.html".
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
EMBL; U07932; AAA21145.1;
DR PIR; I38533; I38531.
DR Genew; HGNC:7138; MLLT6.
DR MIM; 60328;
DR GO; GO:0007048; P oncogenesis; TAS.
DR GO; GO:0006355; P regulation of transcription, DNA-dependent; TAS.
DR InterPro; IPR001965; Znf_PHD.
DR Pfam; PF00628; PHD.1.
DR SMART; SM00249; PHD.2.
DR PROSITE; PS01359; ZF_PHD_1; 1.
DR PROSITE; PS00016; ZF_PHD_2; 2.
KW Nuclear protein; Zinc-finger; Repeat; Chromosomal translocation;
KW Proto-oncogene.
FT ZN_FING 5 57 PHD-TYPE 1.
FT ZN_FING 51 68 C4-TYPE.
FT ZN_FING 111 180 PHD-TYPE 2.
FT DOMAIN 190 211 GLY/SER-RICH.
FT DOMAIN 275 282 POLY-SER.
FT DOMAIN 326 338 POLY-SER.
FT DOMAIN 729 764 LEUCINE-ZIPPER.

```



```

FT DOMAIN 822 829 POLY-SER.
FT DOMAIN 834 862 PRO-RICH.
FT DOMAIN 935 984 GLN-RICH.
FT DOMAIN 1040 1051 POLY-RICH.
FT DOMAIN 1069 1080 GLY-RICH.
FT SITE 551
SQ SEQUENCE 1093 AA; 112021 MW; 753426035736 CRC64;

Query Match
Best Local Similarity 21.4%; Pred. No. 0.5;
Matches 108; Conservative 48; Mismatches 161; Indels 161; Gaps 30;

QY 1: KLAQQT---PKSLRPFTE/GEISLQYSISFLKQLIAS- - - - KLCSSVP 56
Db 590 RLRSPTSLPSSASISTQV---SLAGSTFSPSTHIFGTGPMGAVNPLSQAERSHT 647
QY 57 DPEDLLIYCKRLKDDQTLDFYGIQGSTVIV- - - - -LRKSMPEFDQKEE 102
Db 648 EPDLEDCSPKRGTSPOESLS- - - - -SMSPASLPALEPQCTASPCGGQGLDPAAPGTINWQ 705
QY 103 PVDK- - - - -VAAMREFRVLHTALSSSSSYREAFKMLSNKESLDQIIVATGCLSSDP 155
Db 706 LSEKGGGCEAGNIVEMKADHAKENRLOELTSLSTAKKERQILNV- - - - - 755
QY 136 ALGVLOOKLFSVFALPMNLTIVPAHPALVNAIVLVLSVAGSAPMPPTD- - -SSSKM 212
Db 756 - - - - -CLSVPEFALPAALPAAGPPVPGVGLFPQALSSSS-LSKSP 797
QY 213 PSSSYRMPGGLPEGISSDEDFHNTSTSSSTPSRPSALSYSGAAPPPLITSEL 472
Db 798 POKSSIG - - - - -LNSLSTSSD - - - - -FSGSPSSSSSEFHTEFLPQLQSP 844
QY 273 AT-ALALASTPSSSHTP- - - - -TPTQGHSTSTSPSSGYGCTPTPTPLFSCA-QHA 325
Db 845 ATTLALPAGAPAPPPQPGONGLGAFRAAG - - - - -LAVEMAEGLGLAGGGLPUNGLLG 902
QY 326 L - - - - -CASQPSLO - - - - -SCHQFLQCGPMLQDDELSL 357
Db 903 LNGAAANPAASLQAGGAPTQLDPTGNSLTELQGHFLQCGGCGQCGQ - - - - -QLLA 955
QY 358 RPKRPP - - - - -VATSPKPNSSSILEE- - - - -PHELPA 394
FT 956 SPQTPERTQTVYQMIQI-CQKRELQGLQVAGASQIPMAQLIAGSITLMSASTPGLPT 1015
QY 395 SPEPPANCRGCPWEALMAYVPSL 498
Db 1016 ASAPG - - - - -LEFANAL VAPG 542

RESULT 5
YALG SCHPO STANDARD; PRT: 354 AA.
AC Q12169.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein C26A3.16 in chromosome 1.
GN SPAC26A3.16.
OS Schizosaccharomyces pombe (Fission Yeast)
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes.
CC Schizosaccharomycetales; Schizosaccharomycetaceae;
CC Schizosaccharomycetes.
OX NCBI_TAXID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21849401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream V.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell D., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris E., Hida G., J. J. Hodgson G.

```

```

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitz E.,
RA Rutherford K., Patter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds K., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voickart G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vansteets E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Putnelle B.,
RA Goffeau A., Cadieu E., Dreane S., Gloux S., Lelaure V., Motier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Dominguez A., Revuelta J.L., Moreno S., Armatrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Messy D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-890(2002)
CC - - - SIMILARITY: TO YEAST DSK2
CC - - - SIMILARITY: Contains 1 ubiquitin-like domain.
CC - - - SIMILARITY: Contains 1 UBA domain.
CC - - - SIMILARITY: Contains 1 UBA domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
CC EMBL: 269240; GAZ97239.1;
DR PIR: T33404; T38704;
DR GeneDB SPombe; SIAC26A3.16;
DR InterPro: IPR006436; ST1.
DR InterPro: IPR005449; UBA domain.
DR InterPro: IPR006426; Ubiquitin.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF00245; Ubiquitin; 1.
DR SMART: SM00727; ST1; 2.
DR SMART: SM00165; UBA; 1.
DR SMART: SM00213; UBO; 1.
DR PROSITE: PS00330; UBA; 1.
DR PROSITE: PS00953; UBIQUITIN_2; 1.
KW Hypothetical protein.
FT DOMAIN 363 353 UBIQUITIN-LIKE.
FT DOMAIN 363 353 UBA.
SQ SEQUENCE 354 AA; 36819 MW; 1A99E2D97E73A931 CRC64;

Query Match
Best Local Similarity 6.5%; Score 138.5; DB 1; Length 354;
Matches 90; Conservative 52; Mismatches 120; Indels 165; Gaps 15;

QY 39 SISLKLQIAGKLOESYDP--PELDLYCGKRLKDDQTLDFYGIQGSTVHVLKSPR 96
Db 25 SVLALKEAIA-----PVADIEKRCRLVAGRVLKDEESLTKYTKDQGHSLHVLTKQ 79
QY 97 P-----DQKP-----EPVDKYAAMREFRVLHTALHSSSY----- 126
Db 80 PAAATNVSTORTQCPVNIQAGGANNPLANLTSAR-YAGFNI MPMSASMGFPNPVP 138
QY 127 -----REAVFMLSNKSISLDQIIVATPGL-SSDPIALGVLOKDLFSVFA 170
Db 139 PSTELANLKNPNVQSSINEMESNPQMLDMINSPHLRNAPPYVROMQOSPEFRANT 198
QY 171 DPNMLTLVPAHPALVNAIVLVLSVAGSAPMPGTDSRSSSPSSSYSDMPGGLFEGLS 230
Db 199 DPQVYFQMAQLHQQM-----GAAGIDPM-----SLMG-----GLG 229
QY 23: DDEDFHNTSTPSSTPSRPSALSYSGAAPPPLITQSELATALALASTPSSHTPT 290

```

DE 230 G-----ACLGSGGAG-----240

QY 291 PTOGHSSGTSMSGVSGTPTITLDSFQALCHALQ-----A 328

DB 241 ---LGGFGANNATAGIAGAPVDCTAANNTIONLXNLGGAGGAGLGDACLAGLGGAG 297

QY 329 SCQPSLQSQWQP-----CLOQLPDMQLQDELSRRCRPVPTTSKQWSSSLLEE 378

DB 298 ASPAPAOCTRPPERYAQLSQNEMGFVFFERNVQALRPS-IGNVQGA:ESLJSD 153

RESULT 6

EBN6_EBV

AC P03224; STANDARD; PRT: 992 AA

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-AUG-1988 (Rel. 08, Last sequence update)

DT 01-FEB-1994 (Rel. 28, Last annotation update)

DE EBNA-6 nuclear protein (EBNA-3C) EBNA-4B

GN BFR3-BERF4

OS Epstein-Barr virus (strain 95-8) (Human herpesvirus 4)

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae

OC Gammaherpesvirinae; Lymphocryptovirus

OX NCBI_TaxID=10377

RN (1)_TaxID=10377

RP SEQUENCE FROM N.A.

RX MEDLINE=84277667; PubMed=6237145

RA Baer R., Bankier A.T., Biggin M.C., Deininger P.L., Farrell P.G., Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.J., Segall C.E., Tuffnell P.S., Barrett B.G.

RT "DNA sequence and expression of the 95-8 Epstein-Barr virus genome"

RL Nature 310:207-211 (1984)

RN [2]

RE CHARACTERIZATION

RX MEDLINE=8875572; PubMed=2831394

RA Petri L., Sample J., Wang F., Kieff E.

RT "A fifth Epstein-Barr virus nuclear protein (EBNA5) is expressed in latently infected growth-transformed lymphocytes"

RL J. Virol. 62:1330-1338 (1988)

RN [3]

RP SUBCELLULAR LOCATION

RX MEDLINE=90266473; PubMed=2161150

RA Petri L., Sample C., Kieff E.

RT "Subnuclear localization and phosphorylation of Epstein-Barr virus latent infection nuclear proteins"

RL Virology 176:561-574 (1990)

CC (1)_FUNCTION: INVOLVED IN LATENT VIRUS

CC (1)_SUBCELLULAR LOCATION: NUCLEAR; ANCHORING COMPONENT OF THE PORE COMPLEX

CC (1)_SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, EBNA 5, AND EBNA 6

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch)

EMBL: V01555; CAA24859.1

DR Pfam: PF05009; EBNA-3; 1

KW Nuclear protein; Repeat

FT DOMAIN 74 80

FT POLY-ARG

FT DOMAIN 551 610

FT 10 X 5 AA TANDDEM REPEAT

FT DOMAIN 741 779

FT 3 X 11 AA TANDDEM REPEAT

SQ SEQUENCE 992 AA; 109129 MW; 308EAB9B2515ED54 CWC47

Query Match 6.5%; Score 138; DB 1; Length 992

Best Local Similarity 22.5%; Pred No. 0.55;

Matches 71; Conservative 35; Mismatches 116; Indels 4; Gaps 15;

QY 177 TLVFAHPALVNAIVLHLS---VAGSAPMPGT---205

DE 461 TVEPAHTPVENPMWILLQPPVPVPVVKPTFPSPSRHRRACVYDMLNIEVDVETIE 520

QY 206 DSSSRMPSSSYRDMFGGFLFEGLSQ-----DEDDFHPNTRSTPSSSRPASLQ 257

DB 521 DSSVSCPKPKPKKHODGFGQSRGRKRAAPPTVSPDGTGPPAVGPPAAGPPAAGPAAG 580

QY 258 YSG-----AAGPRITSELTALALASTPSSSHTPTGTO-----294

DB 581 PPAAGPPAAGP-PAAGPRILAPLS-AGPPAAGPHIVTPPSARPRIVAPVVMFMRERQL 638

QY 295 GHSSGTSNVS--GVQSGTPTINDLSQALQHALQSCQPSL-QSQHQPQLQL--RDMG- 349

DB 639 POSTGRKPPQFWEXRAGREITQ--MQCEPSSHLQSATQTPRPSWAPSVCAVSMDACK 696

QY 350 ---IQDELSLRCPRPVGTSKQWSSSLLEEFH-----ELPASPEPPASCRGY-- 395

DB 697 AQPIESSHSSMSPTQPIGHEQP-----RYEDCPAPLDLSLHPDVAAPQAPQYQYQE 752

QY 396 -----CPWEALMKVPP 406

DB 753 PPAPQAPYQYQYQEP 768

RESULT 7

N153_RAT

ID N153_RAT STANDARD; PRT: 1468 AA

AC P49791

DT 01-OCT-1996 (Rel. 14, Created)

DT 01-OCT-1996 (Rel. 14, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Nuclear pore complex protein Nup153 (Nucleoporin Nup153) (153 kDa nucleoporin)

GN NUP153

OS Rattus norvegicus (Rat)

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus

OX NCBI_TaxID=10116

RN (1)_TaxID=10116

RP SEQUENCE FROM N.A. AND SEQUENCE OF 602-613; 622-645 AND 971-993

RX STRAIN=Buffalo; T-SSUE=Liver

RX MEDLINE=93137325; PubMed=8422679

RA Sukegawa J., Blobel G.

RT "A nuclear pore complex protein that contains zinc finger motifs, binds DNA, and faces the nucleoplasm"

RL Cell 72:29-38 (1993)

CC (1)_FUNCTION: POSSIBLE DNA-BINDING SUBUNIT OF THE NUCLEAR PORE COMPLEX (NPC). THE REPEAT-CONTAINING DOMAIN MAY BE INVOLVED IN ANCHORING COMPONENTS OF THE PORE COMPLEX TO THE PORE MEMBRANE.

CC (1)_SUBUNIT: INTERACTS WITH SERP2 (BY SIMILARITY)

CC (1)_SUBCELLULAR LOCATION: Nuclear pore complex; Located to the terminal rim structure of the nucleoplasmic cage.

CC (1)_DOMAIN: Contains F-X-F-G repeats.

CC (1)_PTM: Phosphorylated in interphase, hyperphosphorylated during mitosis. May play a role in the reversible disassembly of the nuclear pore complex during mitosis.

CC (1)_SIMILARITY: THE REPEAT REGION COMPOSED OF PENTAPEPTIDE REPEATS SEPARATED BY SER/THR-RICH DOMAINS IS SIMILAR TO THAT OF YEAST NUP1, NUP2, POX 121 AND MAMMALIAN P62.

CC (1)_SIMILARITY: Contains 4 RANBP2-type zinc fingers.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch)

EMBL: L06821; NOT ANNOTATED_CDS

DR Pfam: A44345; A44345

DR InterPro: IPR001476; Znf_RanGDP

DR Pfam: PF03093; Nucleoporin FG; 9

DR Pfam: PF00641; Znf_RanBP; 4

DR SMART: SM00547; Znf_PBZ; 4

DR PROSITE: P501359; ZF_RANBP2 1; 4.
DR PROSITE: P501359; ZF_RANBP2 2; 4.
KW Nuclear protein; Transpos; Repeat; Zinc-finger; DNA-binding.
KW Phosphorylation.
FT DOMAIN 4 15 GLY-RICH.
FT DOMAIN 442 446 POLY-GLY.
FT ZN_FING 657 687 RANBP2 TYPE 1.
FT ZN_FING 721 750 RANBP2 TYPE 4.
FT ZN_FING 789 818 RANBP2 TYPE 1.
FT ZN_FING 846 875 RANBP2 TYPE 4.
SQ SEQUENCE 1468 AA; 152423 MW; 64596970356AAVC Q164;
Query Match 6.4%; Score 136; DB 1; Length 1468.
Best Local Similarity 23.4%; Pred. No. 13.
Matches 90; Conservative 45; Mismatches 135; Indels 115; Gaps 17;
QY 31 GEYSGLG---GYSISPLKQLIAGLQSVDFELIULIYCKRL --- --- --- --- --- KQD 73
DB 1073 GGFTEGVDSAAISSPDMFVIGRTERQGPVSTSEVFGKACNDEPKQPVFSPGNSE 1132
QY 74 QTLDFYGIQPGSTVHLRYSWPEFDQPE--- --- --- --- --- PVKVAANPERVRLTAHS 122
DB 1133 QTKDESSRPTFSVAKVSKVKESTQAKATFNGQNTWTTTGGAKAFSFLSSSSS 1192
QY 123 SS-----SYREAVPKYLSNKHSEDLIVATPGSSDRIALGVCKKQDSVRAEPNMCATL 178
DB 1193 SSTPATSSSSAIGS STSSSPFVAEFVFCASNPVS--- --- --- --- --- SSAPNSASSTSS 1242
QY 179 VP---AHPLALNAIVLVHVSACAP--MEGTDSSSPKSSPSSVYFMPKPSFEELS 230
DB 1243 QELLFPDQCKPATTS --- --- --- --- --- STAAARPPVFGTQSSNSTVSS--- SEFTGAT 1288
QY 231 DQEDDEHPNTRSPSS STSSSPA--- --- --- --- --- SLYVS-GAARHRTQSELAT 274
DB 1289 --- --- --- --- --- TSSSGSGFVPTGTHKPSASAPAFAGANQTFTFGSQGASVNPSPFSIS 1339
QY 275 ALALAS --- --- --- --- --- TPSSSHIT --- --- --- --- --- TPSTQHSSTPMSQVSGSTPTINIF 318
DB 1340 STAFESAGSQVPPPTGIVSSSSQGVVQDPSGAPSGSTANASSVCFPSSSTINIF 1399
QY 319 SQ --- --- --- --- --- ALQALQASTGSS 333
DB 1400 TNNPNSQVFTFGASPTFAAANAGS 1424
RESULT 8
TCNA TRYCR STANDARD: P230 1101 1401
AC P23253;
DE 01-NOV-1991 (PUB. 20, 1991) 1101 1401
UT 01-NOV-1991 (PUB. 20, 1991) 1101 1401
DT 28-FEB-2003 (PUB. 41, 1991) 1101 1401
DE Stalidase (EC 3.2.1.14) (Neuraminidase) (NA 14) (Enzyme) (Enzyme)
TCNA
CS Trypanosoma cruzi.
CC Eukaryota; Eucelozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
CX NCBI_TaxID=6693;
FN [1]
RP SEQUENCE FROM N.A.
PC STRAIN=Silvio X 10/4;
RX MEDLINE=9127609; PubMed=111561.
SA Pereira M.E.A., Mejia J.S., Ochoa Gutierrez E., Matos evila B.
RA Prioli R.P.;
RT "The Trypanosoma cruzi: neuraminidase contains several sites similar to bacterial neuraminidases, two repeats of the low density lipoprotein receptor, and type III modules of fibronectin".
PL J. Exp. Med. 174:179-191(1991).
FN [2]
RP SUBCELLULAR LOCATION
PX MEDLINE=9134547; PubMed=1896773.
PA Prioli R.P., Mejia J.S., Agui T., Akawa M., Pereira M.E.A.
RC "Trypanosoma cruzi: localization of neuraminidase on the surface of trypomastigotes".
PT

QY 355 LSLRRCRPPVGTGKQPMSSSLDEHPHELPAPEPPASCRG 394
 DB 316 ---RP-EPPAGMASPPGGRS-----PHRLPKDQRKSLALAG 346

RESULT 12

NR41 CANFA STANDARD: PRT: 598 AA.
 ID NR41 CANFA PRT: 598 AA.
 AC P51656;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 30-MAY-2003 (Rel. 35, Last annotation update)
 DE Orphan nuclear receptor HMR (Orphan nuclear receptor; NCBI: NR4A1 OR HMR.
 GN NR4A1 OR HMR.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 CX NCBI_TaxID=9615;
 RN 11;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97050607; PubMed=8895335;
 RA Pichon B., Jimenez-Cervantes C., Pirson I., Maenhaut C.,
 RA Christophe D.,
 RT "Induction of nerve growth factor-induced gene-B (NGFI-B) as an early
 RT event in the cyclic adenosine monophosphate response of doc
 RT thymocytes in primary culture."
 RL Endocrinology 137:4691-4698(1996).
 CC -!- FUNCTION: Orphan nuclear receptor.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NP4
 CC subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>
 CC or send an email to license@isb.sib.ch).

DR EMBL: X97226; CAA65863.1;
 DR HSPB: P19793; 2NLL;
 DR TRANSFAC: T04739;
 DR InterPro: IPR000536; Hormone_recep;
 DR InterPro: IPR001723; Steroid_recep;
 DR InterPro: IPR001628; Znf_C4steroid
 DR Pfam: PF00104; hormone_rec; 1
 DR Pfam: PF00105; znf_C4; 1
 DR PRINTS: PR00398; SDRCHRMONE;
 DR PRINTS: PR00047; STROCFINER;
 DR PRODOM: PD000035; Znf_C4ster; 1;
 DR SMART: SM00433; HCL1; 1
 DR SMART: SM00359; Znf_C4; 1
 DR PROSITE: PS00031; NUCLEAR_RECEPTO;
 KW Receptor; Transcription regulation; DNA binding; Nuclear protein;
 KW Zinc-finger; Phosphorylation;
 FT DNA_BIND 267 332 NUCLEAR RECEPTOR TYPE
 FT ZN_FING 267 287 C4-TYPE
 FT ZN_FING 303 327 C4-TYPE
 FT DOMAIN 409 459 LIGAND BINDING POTENTIAL
 FT DOMAIN 82 92 POLY-SER
 FT DOMAIN 583 586 POLY-PRO
 FT MOD_RES 341 341 PHOSPHORYLATION (BY PKA; BY SIMILARITY)
 FT MOD_RES 351 351 PHOSPHORYLATION (BY PKA; BY SIMILARITY)
 SQ SEQUENCE 598 AA; 64424 MW; 772563A8EC4F906 CRC64;

Query Match 6.0%; Score 129; DB 1; Length 598;
 Best Local Similarity 25.6%; Pred.No. 1;
 Matches 72; Conservative 26; Mismatches 119; Indels 64; Gaps 10;

QY 147 TPG-----LSSDPIALGVLO-CKDLFSVFAQDMCTILVPAHPALVNAIV --LVLHSA 197

DB 14 SPGRPHLASDPLTPELSKPTMDLASPEAAPT-APTALPSPSTMDGTGTFDFTLYOLP 72
 QY 198 GSAPMPGTSSSRSPSS-----YRDMFG---GFLFEGLSDEDDFH 237
 DB 73 GTACPCSSASSASSSTSSSATSPASAFKPEDFQVYCCYGPGLSDPLDELSSGSDYY 132
 QY 238 PNTRSTPSSSTPSPSRPASLG-YSGAAGRPRTIQSELATALALASTPSSSHPTTPGTQGH 296
 DB 133 GSPCSAPSPSTPSPFPQPLSPWDGSPFPSPFSQTYEGLRAWTEQLPKASGHPOPPAFFSF 192
 QY 297 SSGTSPYSSSWQSGTPT:INDLSALQHALCASQPSLQSQMOPQLQOLRDMG:ODDELS 356
 DB 193 SPPTGFSISLACSPKLK-----FPSQATCOLGERESYS:STAPPG 232
 QY 357 LRPCRPVGTGSKQPMSSSLDEHPHELPAPEPPASCRGYCP 397
 DB 233 LAPTSFH-----LOGPGLMDA-PVPSAKARSGAP 260

RESULT 13

RA23 YEAST
 ID RA23 YEAST STANDARD: PRT: 398 AA.
 AC P32628;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE UV excision repair protein RAD23.
 GN RAD23 OR YEL037C OR SYGP-ORF29.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.
 CX NCBI_TaxID=4932;
 RN 11;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94067136; PubMed=8246991;
 RA Watkins J.F., Surg P., Prakash S.,
 RT "The Saccharomyces cerevisiae DNA repair gene RAD23 encodes a nuclear
 RT protein containing a ubiquitin-like domain required for biolog.cal
 RT function."
 RL Mol. Cell. Biol. 13:7757-7765(1993).
 RN 11;
 RP SEQUENCE FROM N.A.
 RC STRAIN=B-6441;
 RX MEDLINE=94016559; PubMed=8411151;
 RA Melnick L., Sherman F.,
 RT "The gene clusters ARC and COR on chromosomes 5 and 10, respectively,
 RT of Saccharomyces cerevisiae share a common ancestry."
 RL J. Mol. Biol. 233:382-388(1993).
 RN 11;
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288c / AD-1;
 RX PubMed=9169868;
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartrell G.,
 RA Huntick-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,
 RA Lin D., Moseedale D., Nakahara K., Namath A., Norgren R., Oetner P.,
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.,
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.;"
 RL Nature 387:78-81(1997)

CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -!- SIMILARITY: Contains 1 ubiquitin-like domain.
 CC -!- SIMILARITY: Contains 2 UBA domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.sib.ch/announce/>
 CC or send an email to license@isb.sib.ch).

DR EMBL: L25428; AAA16070.1; 1;
 DR EMBL: L22172; AAA34935.1; 1;
 DR EMBL: L22173; AAA34938.1; 1;
 DR EMBL: S65964; AAD13972.1; 1;
 DR EMBL: S66117; AAB28441.1; 1;
 DR EMBL: U18779; AAB65005.1; 1;
 DR PIR: S50507; S50507;
 DR HSSP: P54725; IDV0;
 CP SGD: S0000763; Rad23;
 DR GO: GO:000011; Cnucleotide excision repair; factor; 2 comp. ex. 1; A.
 DR GO: GO:000118; Crepairasome; 1; A.
 DR GO: GO:0001684; Fdamaged DNA binding activity; 1; A.
 DR InterPro: IPR034806; Rad23;
 DR InterPro: IPR034806; Rad23;
 DR InterPro: IPR034806; Rad23;
 DR InterPro: IPR034806; Rad23;
 DR InterPro: IPR034806; Rad23;
 DR Pfam: PF00627; UBA; 2;
 DR Pfam: PF00243; Ubiquitin; 1;
 DR SMART: SM00727; STI1; 1;
 DR SMART: SM00155; UBA; 2;
 DR SMART: SM00213; UBQ; 1;
 DR TIGRfam: TIGR00601; rad23; 1;
 DR PROSITE: PS03030; UBA; 2;
 DR PROSITE: PS00853; Ubiquitin 2; 1;
 KW DNA damage; DNA repair; Nuclear protein; Repeat.
 FT DOMAIN 1 77 Ubiquitin-like
 FT DOMAIN 146 186 UBA 1
 FT DOMAIN 355 395 UBA 2
 FT CONFLICT 277 277 A -> S (in Ref. 2)
 SQ SEQUENCE 398 AA; 42366 MW; B3FC416DAB68B833 CRC64;
 Query Match 6.0%; Score 227; DB 1; Length 398;
 Best Local Similarity 24.8%; Pred. No. 8;
 Matches 57; Conservative 28; Mismatches 125; Indels 43; Gaps 8;
 QY 118 TALHSSSYREAVFKML-SNESLQGLVATPGL-SSDPALGVLOKCLFSVHADPNYL 175
 DB 29 TKLAQS-SCESQ-KLIVESVVLGKATVSCGKQKQGVVYSQKSKTKVTEP- 85
 QY 176 DTLPAHPALVNA-VLVLSVAGAPVVGTSSTSSSYSPSSSYRPMQGFIFESLEDDQD 235
 DB 86 -----PAAE----- -SATTGRENTEASPTSDASAPATAPAG-SQCEE 126
 QY 236 FHPNTRTPSSSTPS-----SPPASGVGSGAGPFF- -TCELLATALLAS 280
 DB 127 QTATTERTESASTPGVVGVTGNETIERIMGVQREVEERAEALFAFNHRAVEYLK 186
 QY 281 TPESSTHTPTGTGCHSSSTSPMSEVGVSGTPTTNVLFSQLALCALQAS 140
 DB 187 IPE NIPETPTGCTAAMAF-FTSTLFAELPGLTFAACGLTFAAG 140
 RESULT 14
 ID ICP4 HSVEB STANDARD; PRT: 1487 AA
 AC P28925;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 01-FEB-1994 (Rel. 28, Last annotation update)
 DE Trans-acting transcriptional protein ICP4 (155 kDa immediate early protein).
 GN IE OR 64.
 OS Equine herpesvirus type 1 (strain Abp) (Herpesviridae; dsDNA viruses, no RNA stage; Herpesviridae; Alphaherpesvirinae; Varicelloviruses).
 OX NCBI_TaxID=31520;
 RN 111
 RP SEQUENCE FROM N.A.
 RA MEDJ:NM=92295566; PubMed=1118606;
 RA Telford E.A.R.; Watson M.S.; McBrade K.; Davison A.C.;
 RT "The DNA sequence of equine herpesvirus-1";
 RL Virology 189:304-316(1992).
 CC -!- FUNCTION: THIS IE PROTEIN IS A MULTIFUNCTIONAL PROTEIN CAPABLE

CC OF MIGRATING TO THE NUCLEUS, BINDING TO DNA, TRANS-ACTIVATING
 CC SUBCELLULAR LOCATION: NUCLEUS OF INFECTED CELLS.
 CC -!- PTM: A LONG STRETCH OF SERINE RESIDUES MAY BE A MAJOR SITE OF
 CC PHOSPHORYLATION.
 CC -!- SIMILARITY: BELONGS TO THE ICP4/IE140/IE180 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 DR EMBL: M86664; AAB2515.1; 1;
 DR EMBL: M86664; AAB2515.1; 1;
 DR PIR: A36802; EDB376;
 DR TRANSFAC: T0C401;
 DR InterPro: IPR005405; Herpes_ICP4_C.
 DR InterPro: IPR005406; Herpes_ICP4_N.
 DR Pfam: PF03585; Herpes_ICP4_C; 1;
 DR Pfam: PF03584; Herpes_ICP4_N; 1;
 KW Early protein; Transcription regulation; Trans-acting factor;
 KW DNA-binding; Phosphorylation; Nuclear protein.
 FT DOMAIN 181 213 SER-RICH.
 FT DOMAIN 922 931 ARG/LYS-RICH (BASIC).
 SQ SEQUENCE 1487 AA; 154868 MW; 1D4397838D03680D CRC64;
 Query Match 6.0%; Score 127; DB 1; Length 1487;
 Best Local Similarity 21.2%; Pred. No. 4;
 Matches 85; Conservative 41; Mismatches 138; Indels 136; Gaps 18;
 QY 48 AGKLGESVPELIDLIYCGRKLDQTLDFYGIQ--PGSTVHVL---RKSMPER----- 97
 DB 30 ASAAEGFTCPAAPDLLYGSQNM-----FGVDDAPLSTPAVIPPSTPPEPGCKA 81
 QY 98 -----DQKPEPVKVAAMREFRVLHTALHSSSYREAVFKMLS 135
 DB 82 KRSPSAAGSGGPTTAAQAAPASPAAPGLAAM--LKMVHSSVAPGNRR----- 131
 QY 136 NKEISDQGLVAT-----PGLSSDPALGVLOKCLFSVHADPNMLDVLPAHPALVNAI 189
 DB 132 -----ATGSSSPGGDAADPVAL-----DSDTETCPGSPQ-----PEPPSA--- 168
 QY 190 VLVLSVAGAPVVGTSSTSSSYSPSSSYRPMQGFIFEGL-----SDDEDHPNPT 240
 DB 169 -----SPGSGSPAPVRS;SISSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 223
 QY 241 ESTP-----SNETPSSRPASGVSG--AAGRPPTQSELATALLASTPES--SHTPTPGTQ 294
 DB 224 ESTTPPESVNAAKTSAAGSPGSSGGDP-----AAGAATPKSCRSGAASPCAP 274
 QY 295 GHSSGTSVSSGVQSGTPTTNLFSCALQHALCASQPSLSQSQWQPQLQQ-RDMGIQDDE 354
 DB 275 APASATAFESRPGGLFPGARILEYLEGVTEANLAKTLE----- 315
 QY 355 LSUPPCRETVGTSKQWSSSLLEPHELPASPEPPASCRG 394
 DB 316 ---RP-EPNAGMASPPGRS-----PHRLPKDQRPKXSALAG 346
 RESULT 15
 ID SCR1 SCHPO STANDARD; PRT: 565 AA.
 AC O14335;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA-binding protein scr1.
 GN SCR1 OR SPBC1D7.02C.
 OS Schizosaccharomyces pombe (fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CorpuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 06:35:08 ; Search time 41 Seconds
(without alignments)
2580 536 Million cell updates/sec

Title: US-09-735-251-4
Percent score: 2.33
Sequence: 1 MSLSDWHLAVKLADQFLPRK SSGVCPALMKVPPSPRY 419

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.6

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 53525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0

Maximum Match 100

Listing first 45 summaries

Database : SP-REMBL231*

1: sp_archaea*

2: sp_bacteria*

3: sp_fungi*

4: sp_human*

5: sp_invertebrate*

6: sp_mammal*

7: sp_thc*

8: sp_organelle*

9: sp_phage*

10: sp_plant*

11: sp_rodent*

12: sp_virus*

13: sp_vertebrate*

14: sp_unclassified*

15: sp_virus*

16: sp_bacteriap*

17: sp_archaeap*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total 1000-Base-Dictionary.

SUMMARY

Result NO.	Score	Query Match	Length	DB	IE	Description
1	1834	86.0	380	4	Q96103	Q96103 homo sapien
2	1824	85.5	380	4	Q96S82	Q96S82 homo sapien
3	1804	84.6	380	11	Q91W67	Q91W67 mus musculus
4	1780	83.5	380	11	Q9D7P5	Q9D7P5 mus musculus
5	189.5	8.8	554	11	Q8K141	Q8K141 mus musculus
6	162.5	8.6	548	10	Q8LSP9	Q8LSP9 oryza sativ
7	161.5	8.5	554	11	Q8CB35	Q8CB35 mus musculus
8	178.5	8.4	592	11	Q8R317	Q8R317 mus musculus
9	177	8.3	555	11	Q8C7T4	Q8C7T4 mus musculus
10	175	8.2	523	5	Q9N1F3	Q9N1F3 dictyosteli
11	166.5	7.8	638	11	Q9QZMC	Q9QZMC mus musculus
12	164.5	7.7	538	10	Q9S119	Q9S119 arabidopsis
13	161.5	7.6	547	5	Q9VWD9	Q9VWD9 drosophila
14	158.5	7.4	582	11	Q9JJP9	Q9JJP9 rattus norv
15	152.5	7.1	624	4	Q9JHD9	Q9JHD9 homo sapien
16	152.5	7.1	624	4	Q9HAZ4	Q9HAZ4 homo sapien

17	152.5	7.1	2164	12	Q9DH52	Q9DH52 meleagrid h
18	150	7.0	575	10	Q9FWF5	Q9FWF5 oryza sativ
19	149	7.0	415	5	Q9VHX2	Q9VHX2 drosophila
20	147.5	6.9	595	6	Q9SM59	Q9SM59 bos taurus
21	147	6.9	862	11	Q9JIK1	Q9JIK1 rattus norv
22	144	6.8	582	11	Q9QZM1	Q9QZM1 mus musculus
23	143.5	6.7	551	10	Q94C51	Q94C51 arabidopsis
24	143.5	6.7	551	10	Q9S118	Q9S118 arabidopsis
25	142	6.7	2161	12	Q9E1F1	Q9E1F1 meleagrid h
26	141.5	6.6	589	4	Q9UMX0	Q9UMX0 homo sapien
27	141	6.6	528	6	Q29C71	Q29C71 sus scrofa
28	141	6.6	671	10	Q8O910	Q8O910 arabidopsis
29	141	6.6	830	11	Q9NM38	Q9NM38 rattus norv
30	141	6.6	858	11	Q9NM37	Q9NM37 rattus norv
31	140.5	6.6	589	4	Q9H0T8	Q9H0T8 homo sapien
32	139.5	6.5	589	4	Q8IXS9	Q8IXS9 homo sapien
33	139.5	6.5	778	5	Q8MKG7	Q8MKG7 aedes aegypt
34	138	6.5	992	12	Q69119	Q69119 human herpe
35	138	6.5	1325	5	Q3BKV7	Q3BKV7 leishmania
36	137.5	6.4	259	4	Q8NAMS	Q8NAMS homo sapien
37	137	6.4	831	11	Q8CEJ3	Q8CEJ3 mus musculus
38	136.5	6.4	713	3	Q9P6R7	Q9P6R7 schizosacch
39	136.5	6.4	1758	11	Q8CHN9	Q8CHN9 rattus norv
40	135.5	6.4	873	5	Q9Y076	Q9Y076 leishmania
41	135.5	6.4	2715	5	Q61603	Q61603 drosophila
42	135	6.3	1502	5	Q8IS10	Q8IS10 dictyosteli
43	134.5	6.3	1338	4	Q9BZG0	Q9BZG0 homo sapien
44	134.5	6.3	1338	4	Q96RG4	Q96RG4 homo sapien
45	134.5	6.3	1339	4	Q96RG5	Q96RG5 homo sapien

ALIGNMENTS

RESULT 1
Q96103 PRELIMINARY; PRT; 380 AA.
AC Q96103
DT 01-DEC-2001 (TrEMBL:rel. 19, Created)
DT 01-DEC-2001 (TrEMBL:rel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBL:rel. 23, Last annotation update)
DE Similar to RIKEN CDNA 2300004C15 gene (Hypothetical protein MGCL4421).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Uterus;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBSJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC007913; AAH07913.1; -
DR EMBL; BC030055; AAH30055.1; -
DR EMBL; BC033919; AAH33919.1; -
DR InterPro; IPR001923; SECE domain.
DR InterPro; IPR000626; Ubiqutin.
DR Pfam; PF00627; Uba; 1.
DR PROSITE; PS01067; SECE_SEQ61G; 1.
DR PROSITE; PS50053; UBIQUITIN_2; 1.
KW Hypothetical protein.

```

SQ SEQUENCE 380 AA; 40510 MW; 08649F79597E09A1 CRC64;
Query Match 86.0%; Score 1834; DB 4; Length 380;
Best Local Similarity 130.0%; Pred. No. 3 5e-129;
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
D 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
QY 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
D 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
QY 121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180
D 121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180
QY 181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
D 181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
QY 241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
D 241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
QY 301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358
D 301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358

RESULT 2
ID Q96582 PRELIMINARY; PRT; 380 AA
AC Q96582;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Ubiquitin-like protein S5.32.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catartina; Hominidae; HUP-0;
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RA Zhang W., Li N., Wan T., Cao X.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AY037166; ANK67643.1;
DR InterPro; IPR0019C1; SecE;
DR InterPro; IPR006449; UBA domain;
DR Pfam; PF00627; UBA; 1;
DR Pfam; PF00240; Ubiquitin; 1;
DR SMART; SM00165; UBA; 1;
DR PROSITE; PS01067; SEC6; SEC6IG; 1;
DR PROSITE; PS00053; UBIQUITIN 2; 1;
SQ SEQUENCE 380 AA; 40581 MW; 5AF6A14F81F996C CRC64;

Query Match 85.5%; Score 1824; DB 4; Length 380;
Best Local Similarity 99.4%; Pred. No. 2 4e-128;
Matches 356; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
D 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
QY 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
D 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
QY 121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180

```

```

121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180
181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358
301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358

RESULT 3
Q91M67 PRELIMINARY; PRT; 380 AA
ID Q91M67;
AC Q91M67;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Similar to RIKEN cDNA 2300004C15 gene.
GN 2300004C15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_TaxID=10090;
RN 1;
RP SEQUENCE FROM N.A.
RA Straussberg R.;
RA Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC016456; AAH16456.1;
DR MGD; MG:1916709; 2300004C15RIK.
DR InterPro; IPR0019C1; SecE;
DR InterPro; IPR006449; UBA domain;
DR InterPro; IPR00626; Ubiquitin;
DR Pfam; PF00627; UBA; 1;
DR Pfam; PF00240; Ubiquitin; 1;
DR SMART; SM00165; UBA; 1;
DR SMART; SM00213; UBA; 1;
DR PROSITE; PS01067; SEC6; SEC6IG; 1;
DR PROSITE; PS00053; UBIQUITIN 2; 1;
SQ SEQUENCE 380 AA; 40407 MW; 6DE0834A7CCEARF2 CRC64;

Query Match 84.6%; Score 1804; DB 11; Length 380;
Best Local Similarity 97.9%; Pred. No. 6 1e-127;
Matches 350; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
D 1 MSLSDWHLAVKLAQOPLTPKSIILRLPETELGEYSLGYSISFLKQLIAGKLOESVDPDEL 60
QY 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
D 61 IDLIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPDKPEPVCKVAAMREFVLHTAL 120
QY 121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180
D 121 HSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPALGVLODKLFSVFADPNMLDTLVP 180
QY 181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
D 181 AHPALVNAIVLVHVSAGSAPMPGTDSSSRSPSSSYRDMPGFLFEGLSDDDDDFHNT 240
QY 241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
D 241 RSTPSSSTPSSRPASLGYSAGAPRPITQSELATALALASTPSSSHHTPTPTQGHSSGT 300
301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358
301 SPMSGVSQSGTPTINDLFSQALQHALQASGQPSLOSQWQPQLQQLRDMGIDDELSLR 358

```

RESULT 4

Q9D7P5 PRELIMINARY; PRT; 380 AA.
 AC Q9D7P5;
 DT 01-JUN-2001 (TRENBLREL. 17, Created);
 DT 01-JUN-2001 (TRENBLREL. 17, Last sequence update);
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update);
 DE 2300004C15R1k protein.
 GN 2300004C15R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Tongue;
 RC MEDLINE=21085660; PubMed=1127851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kaeska K., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kouchiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido T., Plesch G., Quackenbush J.,
 Schriml L.M., Staab F., Suzuki R., Torita M., Wagner D., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo V.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi V.,
 Gusticich S., Hill D., Hofmann V., Iwata S., Kariya W., Lee K.H.,
 Lyons P., Marchionni L., Mashima M., Mazzarelli L., Mombaefts P.,
 Notodono P., King B., Ringwald M., Rodriguez L., Sakamoto K.,
 Saeki H., Sato K., Schoenbach K., Seya T., Shibata Y., Storch K.F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz M., Whitaker C., Williams L.,
 Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 Hayashizaki Y.
 RA "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 DR EMBL; AK009025; BAB26033.1;
 DR MGD; MGI:1916709; 2300004C15R1K.
 DR InterPro; IPR001901; SecE.
 DR InterPro; IPR000449; SecE domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR Pfam; PF00240; Ubiquitin; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00237; UBA; 1.
 DR PROSITE; PS01067; SECE; SEC6; 1.
 DR PROSITE; PS50953; UBIQUITIN; 1.
 DR PROSITE; PS50953; UBIQUITIN; 1.
 SQ SEQUENCE 380 AA; 40440 MW; 2300004C15R1K; 2300004C15R1K

Query Match 83.5%; Score 188.5; DB 11; Length 554;
 Best Local Similarity 96.6%; Pred. No. 8.2e-06;
 Matches 146; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 1 MSLSDHVLAVKADQPLTPKSLIPETELGYSILGYSISFLKQLIAGKLOESVDPDEL 60
 DB 1 MSLSDHVLAVKADQPLTPKSLIPETELGYSILGYSISFLKQLIAGKLOESVDPDEL 60
 QY 61 IDLYCGKRLKDDQTLDFYGIQPGSTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 120
 DB 61 IDLYCGKRLKDDQTLDFYGIQPGSTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 120
 QY 121 HSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGKLOESVDPDEL 180
 DB 121 HSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGKLOESVDPDEL 180
 QY 121 HSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGKLOESVDPDEL 180
 DB 121 HSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGKLOESVDPDEL 180
 QY 181 AHPALNVAIVLVHVSAGSAPYPTGDSRSSRSPSSSYRDMEGGLFEGLSDDDEPHNT 240
 DB 181 AHPALNVAIVLVHVSAGSAPYPTGDSRSSRSPSSSYRDMEGGLFEGLSDDDEPHNT 240
 QY 241 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTGTQGHSSGT 300

DB 241 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTGTQGHSSGT 300
 QY 301 SPMSGVSQGTPIITNDLFSQALQASGOPSLSQSQWQPOLQQLRDMGTQDDLSLR 358
 DB 301 SPMSGVSQGTPIITNDLFSQALQASGOPSLSQSQWQPOLQQLRDMGTQDDLSLR 358
 RESULT 5
 Q8K141 PRELIMINARY; PRT; 554 AA.
 ID Q8K141;
 AC Q8K141;
 DT 01-OCT-2002 (TRENBLREL. 22, Created);
 DT 01-OCT-2002 (TRENBLREL. 22, Last sequence update);
 DT 01-MAR-2003 (TRENBLREL. 23, Last annotation update);
 DE Hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Retina;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC028857; AAH28857.1;
 DR InterPro; IPR000449; UBA domain.
 DR InterPro; IPR000626; Ubiquitin.
 DR Pfam; PF00627; UBA; 1.
 DR SMART; SM00165; UBA; 1.
 DR SMART; SM00237; UBA; 1.
 DR PROSITE; PS50953; UBIQUITIN; 2; 1.
 DR PROSITE; PS50953; UBIQUITIN; 2; 1.
 DR Hypothetical protein.
 SQ SEQUENCE 554 AA; 58677 MW; P255C9985A4C029 CRC64;

Query Match 8.8%; Score 188.5; DB 11; Length 554;
 Best Local Similarity 23.0%; Pred. No. 8.2e-06;
 Matches 104; Conservative 76; Mismatches 174; Indels 99; Gaps 21;
 QY 15 QPLTPKSLIPETELGYSILGYSISFLKQLIAGKLOESVDPDELIDLYCGKRLKDD 73
 DB 15 QPLTPKSLIPETELGYSILGYSISFLKQLIAGKLOESVDPDELIDLYCGKRLKDD 73
 QY 24 EPKIMKVTGTPK-EKEEFVAVPSSVQFKEEISKRFKSHI---DQLVLI-FAGKILKDD 79
 DB 24 EPKIMKVTGTPK-EKEEFVAVPSSVQFKEEISKRFKSHI---DQLVLI-FAGKILKDD 79
 QY 74 OTLDFYGIQPGSTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 112
 DB 74 OTLDFYGIQPGSTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 112
 QY 80 DTLSQGHGHDGLTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 137
 DB 80 DTLSQGHGHDGLTVHVLKRSMPERCQKPEPVKVAAMKEEFVLHTAL 137
 QY 113 FRV-----LHTALHSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGVL 160
 DB 113 FRV-----LHTALHSSSYREAVFPMUNKESLDQIIIVATPGLSSDIALGVL 160
 QY 138 FVGGLGGAGLSSSLGNTTFSELSQSQWQPOLQQLRDMGTQDDLSLR 190
 DB 138 FVGGLGGAGLSSSLGNTTFSELSQSQWQPOLQQLRDMGTQDDLSLR 190
 QY 161 QKDLFS-VFADPNKLDIVFAHPALVNA-----IVLVHVSAGSAPYPTGDSRSSMP 213
 DB 161 QKDLFS-VFADPNKLDIVFAHPALVNA-----IVLVHVSAGSAPYPTGDSRSSMP 213
 QY 191 SNPDLMLQINANPOM-QOLIQRNPEISHMLNPNDRQTLLEARNPAMMOEMNRNDRA 249
 DB 191 SNPDLMLQINANPOM-QOLIQRNPEISHMLNPNDRQTLLEARNPAMMOEMNRNDRA 249
 QY 214 SSSYRDMEGGLFEGLSDD-----EDFHPNTRSTPSSSTPSSRPASGLYSGAAG 263
 DB 214 SSSYRDMEGGLFEGLSDD-----EDFHPNTRSTPSSSTPSSRPASGLYSGAAG 263
 QY 250 LSNLESIFGYNAIRMYTDIQEPLNAAQFQGN----PFASLVSS-----SBAEG 299
 DB 250 LSNLESIFGYNAIRMYTDIQEPLNAAQFQGN----PFASLVSS-----SBAEG 299
 QY 264 PRPITQSELATALASTPSSSHTPTGTQGHSSGTSPMSGVSQGTPIITNDLFSQALQ 323
 DB 264 PRPITQSELATALASTPSSSHTPTGTQGHSSGTSPMSGVSQGTPIITNDLFSQALQ 323
 QY 300 TQP-SRTENPDLNFWAPQTSQSPASGTTGTTNTMTSGGTATSTPAQQTSGPSLV 358
 DB 300 TQP-SRTENPDLNFWAPQTSQSPASGTTGTTNTMTSGGTATSTPAQQTSGPSLV 358
 QY 324 HALQAS--GQSLQSQWQ-----POLQQLRDMGTQDDLSLRPCRPV-GTSGQPSVSL 375
 DB 324 HALQAS--GQSLQSQWQ-----POLQQLRDMGTQDDLSLRPCRPV-GTSGQPSVSL 375
 QY 359 PGAGASMTNPGMSLLQOITENPOLMQ-----NMLSAPYRSMQLQSLQNPDLAAQ 410
 DB 359 PGAGASMTNPGMSLLQOITENPOLMQ-----NMLSAPYRSMQLQSLQNPDLAAQ 410
 QY 376 LEEPELPAPEPPACRCYCPWALMKVPPSL 408
 DB 376 LEEPELPAPEPPACRCYCPWALMKVPPSL 408
 QY 411 MNPDTLSMNSPRA-----MQALLIQOGL 436
 DB 411 MNPDTLSMNSPRA-----MQALLIQOGL 436

RESULT 6

```

Q8LSP9
ID Q8LSP9 PRELIMINARY; PRT; 548 AA.
AC Q8LSP9;
DT 01-OCT-2002 (T:EMBLrel. 22, Created);
DT 01-OCT-2002 (T:EMBLrel. 22, Last sequence update);
DT 01-MAR-2003 (T:EMBLrel. 23, Last annotation update);
DE Putative ubiquitin protein.
GN Q8LSP9.1.
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Fagaceae;
OC Eucaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX NCBI_TaxID=39947;
RN [1]
SEQUENCE FROM N.A.
RA King R.A., Yu Y., Soderlund C., Chen M., Kim H.-R., Hambro T.,
RA Sasaki C., Henry D., Oates R., Simmons J.,
RT "Rice Genomic Sequence";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DE EMBL: AC098693; AAC22708.1;
DR Gramene; Q8LSP9;
DR InterPro: IPR006636; STIL;
DR InterPro: IPR000449; UBA domain;
DR InterPro: IPR000626; Ubiquitin;
DR Pfam: PF06627; UBA;
DR PRINTS: PR00348; UBQUITIN;
DR SMART: SM00727; STIL; 4;
DR SMART: SM00165; UBA; 1;
DR SMART: SM00213; UBC; 1;
DR PROSITE: PS02299; UBQUITIN; 1;
DR PROSITE: PS00053; UBQUITIN; 2;
DR PROSITE: PS00053; UBQUITIN; 2;
SQ SEQUENCE 548 AA; 57797 MW; F4C6E75A8FCB5 CR64;

Query Match 8.6%; Score 181.5; DB 10; Length 548;
Best Local Similarity 23.4%; Pred No 2.7e-05;
Matches 104; Conservative 60; Mismatches 173; Indels 107; Gaps 21;

QY 37 GYSIFLQIAGKLOESVPEPELIIYCPKPKDKDTLLYHQPGSTVHLKSNPE 96
DB 35 GLSVAEKAIVAESG--DVPAQC RITYKGFILKDEGLASVAVETHTHVEGAATP 91
QY 97 PQDKPEP-----VQVVAWMEERV-----LFTALHSSSSSYREAVFKM 143
DB 92 PASTAPFAANNVTFAINTATTASNAFAVFGGILHGLSGSASGUGLQFISGLPELSW 151
QY 144 - IVATPGLSST---PGLATDLDLFS VFADPNMLDVLPAHPALVNA----- 160
DB 152 QVQSENTEHPELWNNK VLDVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 160
QY 193 LHSVAGSAMP - VFADPNMLDVLPAHPALVNA-----IVLVHVSAGSAMPPTDSSSRMP 213
DB 211 LQCTVEAARNPELWREMMNTDRAVNSLSS -VFADPNMLDVLPAHPALVNA----- 211
QY 232 -DEDFEHN-----PGLATDLDLFS VFADPNMLDVLPAHPALVNA----- 259
DB 266 AGEGRSSNPFSALLNGNSQACDFPAASTTTTSTTSGSPATNHLNAINITACAA 325
QY 260 -GAAGPRITQSELATALASTPSSSHTPTGCGHSSGTSIPMSGVSGTPTINDLFSQALQ 318
DB 326 QGATRPSPVTNARSATAGLGL -SSTLQGMJUGSGDTNLSQVLC -NPEM 375
QY 319 SQALQHAL---QASGQ-----PSLQSQWQPLQCRKMGICFTELSLRPPPPVGTSKQP 370
DB 376 MQMNMNINPQSMNQLNINPNVAMNESN-TCIREV-FQWLEFVRCGLSPETLQQLIS 433
QY 371 WSSSLPEHPALPASPEPPASCPG 394
DB 434 FOQSLMSQGGQACGAPERTQSGAG 457

RESULT 7
Q8C815
ID Q8C815 PRELIMINARY; PRT; 582 AA.
AC Q8C815;
DT 01-JUN-2002 (T:EMBLrel. 21, Created);
DT 01-JUN-2002 (T:EMBLrel. 21, Last sequence update);
DT 01-MAR-2003 (T:EMBLrel. 23, Last annotation update);
DE Similar to ubiquitin 1.
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.

Query Match 8.5%; Score 181.5; DB 11; Length 554;
Best Local Similarity 22.7%; Pred No 2.7e-05;
Matches 103; Conservative 76; Mismatches 175; Indels 99; Gaps 21;

QY 15 QPLTPKSLRLPETELGEYSL-GYGISIFLQIAGKLOESVPEPELIIYCGRKLKDD 73
DB 24 EPKIMKVTYKTPK-EKEEFAPVENSVOQFKEEISKRFKSHI---DQLVLI FAGKILKDD 79
QY 74 QTLDFYGIQPGSTVHLVRK-----SWPEQDKPEPVDKVAAMRE 112
DB 80 DTLQSRGHGGLTTHLV:KIQNRPONSAQOTNAPGSTVTSSPAPDS--NPTSGSAANS 137
QY 113 FRV-----LFTALHSSSSSYREAVFKM-----LSNKEISLDQIIVATPGLSDDPIALGVL 160
DB 138 FGVCGLGGLGSLSSGCTNTNFSLOSGWQRCGLSNPGMMVQIM-----ENPFVQSM 190
QY 161 QDKDLFS VFADPNMLDVLPAHPALVNA-----IVLVHVSAGSAMPPTDSSSRMP 213
DB 191 SNEDLMRLINANPQV-OQLIQNPESHMLNPNPCIMROTTLEARNPAMQWEMNRNODRA 249
QY 214 SSSRYDMPGF-----JFGLSDD-----EDPHNTRSTPSSSTPSSPASLGYSAGAG 263
DB 250 LSNLES:PSYNAIRMYTDIQEPLNAAACDFGNN---PFASLVSSS-----SSAEG 299
QY 264 PRPTQSELATALASTPSSSHTPTGCGHSSGTSIPMSGVSGTPTINDLFSQALQ 323
DB 300 TOP SSSRYDMPGF-----JFGLSDD-----EDPHNTRSTPSSSTPSSPASLGYSAGAG 358
QY 324 HALQAS -PSLQSQWQ-----PGLQQLRDMGIDDELSLRPPPPV-GTSKQPMSSSL 375
DB 359 PGAGASGNTPGVQSILQQTENPQLMQ-----NYLSAPVWMSMLQSLQNPDLAAQ 410
QY 376 LEEPHLPSPEPPASCPGCPWEALMKVPPSL 408
DB 411 MOKPDTLSVSNPRA-----MQALLQIQQL 436

RESULT 8
Q8R317
ID Q8R317 PRELIMINARY; PRT; 582 AA.
AC Q8R317;
DT 01-JUN-2002 (T:EMBLrel. 21, Created);
DT 01-JUN-2002 (T:EMBLrel. 21, Last sequence update);
DT 01-MAR-2003 (T:EMBLrel. 23, Last annotation update);
DE Similar to ubiquitin 1.
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.

```

Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.

```

RL EMBL: BC026847; AAH26847.1;
DR InterPro: IPR006636; STIL;
DR InterPro: IPR000449; UBA domain;
DR InterPro: IPR000626; Ubiquitin;
DR Pfam: PF00627; UBA; 1;
DR Pfam: PF00240; Ubiquitin; 1;
DR SMART: SM00727; STIL; 4;
DR SMART: SM00165; UBA; 1;
DR PROSITE: PS0053; UB-QUITIN; 2; 1;
DR PROSITE: PS0053; UB-QUITIN; 2; 1;
SQ SEQUENCE 582 AA; 61976 MW; 62253735A; 19.72E CRC64;

Query Match: 8.4%; Score: 178; DB 5; Length: 523;
Best Local Similarity: 23.5%; Pred. NO: 4.9e-05;
Matches: 91; Conservative: 63; Mismatches: 149; Indels: 4; Gaps: 18;

CY 15 QPLTKSLRLRLPTELGYSLSGGYSISFJLKOLIAKGLQESVPDPDELIDLYCGKRLKDDT 75
DB 24 EPKIMVATPK-EKEEFAVPENS-SVQOFKEEISKRFKSHI---DQVLVIFAGKILKDDT 79
CY 74 QTLDFYGIOPGSTVHVLRK-----SWPEPDQKPEPVKVAAMREFR 114
DB 90 DTLSCGHGHDGLTVHVLVKTNRPNQNSAQTNAPGSTVTSPPADS--NPTSUSAANSF 137
CY 113 FRY-----LHTALHSSSYREAVFKV---LSNKESLDQIIIVATPGLSSDPIALGVLD 162
DB 138 FGVGGLAGLSSGLNTNFSELSQSQMQQLSNPEMVQIM-----ENFVQSLN 165
CY 161 QKOLFES--VFADPNMIDTLVPAHPALVNA-----IVLVHVSAGSAPMPGTDSSSRMP 213
DB 191 SNPELMROQIMANPQV-QQLCRNPETSHMLNPNQTLLEARNPAMMOEMMRNODRALS 224
CY 214 SSSYRMPGGF-----LFEGLSDQ-----EDDFHPNTSTPTSSSTPSRSPASLGYSGA 265
DB 250 LSNLESIPGGYNALRRMYTDIOEPMNAAQEQFGN-----PFASLVSSS-----SSA 274
CY 264 PRITQSELATALALASTPSSSHPTPTGQGHSSGTSMPSSGVQSGTPTIINDLFSQALQ 325
DB 300 TQP-SRTENRDELPNPWAPQTSOSSPASGTTGGTTNTWTSGGTATSTPAGQSTSGPSLV 333
CY 324 HALCAS--GQPSLQSQMW-----POLQQ 344
DB 359 PGAGSMFNTPGVQSLLQITENPOLMQ 359

RESULT 9
Q807T4 PRELIMINARY; PRT; 523 AA.
AC Q9NIF3
DT 01-OCT-2003 (TEMBLrel: 15, Created)
DT 01-OCT-2003 (TEMBLrel: 15, Last sequence update)
DT 01-MAR-2003 (TEMBLrel: 23, Last annotation update)
DE Ubiquitin-like fusion protein SonA.
GN SONA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelia; Dictyostelium.
OX NCBI_TaxID=44689;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA Pukatzki S., Ennis H.L., Kessin R.H.;
RT "Establishing a genetic interaction between a ubiquitin-like protein
RT and ubiquitin-mediated proteolysis in Dictyostelium discoideum."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK04118; AAF4453.1;
DR HSSB: P02249; 1;
DR InterPro: IPR001614; STIL;
DR InterPro: IPR005447; UBA domain;
DR InterPro: IPR003626; Ubiquitin;
DR Pfam: PF00240; Ubiquitin; 1;
DR PRINTS: PR00148; UBQUITIN;
DR SMART: SM00727; STIL; 4;
DR SMART: SM00165; UBA; 1;
DR SMART: SM00213; UBQ; 1;
DR PROSITE: PS0053; UBQUITIN; 2; 1;
DR PROSITE: PS0053; UBQUITIN; 2; 1;
SQ SEQUENCE 523 AA; 56671 MW; 62F5C127761A5133 CRC64;

Query Match: 8.2%; Score: 175; DB 5; Length: 523;
Best Local Similarity: 22.9%; Pred. NO: 7.8e-05;
Matches: 93; Conservative: 63; Mismatches: 133; Indels: 118; Gaps: 19;

CY 37 GYSISFLKQIAGKIQESVP-DPELIDLYCGKRLKDDTLDYGIOPGSTVHVLRKSWP 95
DB 21 GITVADEKKVATKC--SIPADQRI--IYSGRLKHQOTLDEIKIQGHTVHVLRKGAAP 76
CY 96 EPDQKPEPVK-----VAAMREFRVLTALHSSSYR 127
DB 77 PP---PPVFEQQVTPSTNTQPGIPGVQPNKINDMNNPFIQEMFNSRMMDSLINDPDI 133

```



```

RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carnacci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Kallin-Neumann G., Kawai J., Kim C., Lam B., Lin C.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Saito M., Seki M., Shirai P., Southwick A., Shinzaki K.,
RA Davis R.W., Ecker J.R., Theologis A.,
RT "Arabidopsis Open Reading Frame (ORF) Clones."
RL Submitted (MAY-2002) to the EMBL/GenBank/CDR databases.
DR EMBL: AC007127; AA225137.2;
DR EMBL: AF362159; AA25869.1;
DR EMBL: AY042828; AA68768.1;
DR EMBL: AY081450; AA10012.1;
DR EMBL: AY113886; AA44934.1;
DR InterPro: IPR000443; UBA domain.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF02240; Ubiquitin; 1.
DR PROSITE: PS00053; Ubiquitin-N2; 1.
SQ SEQUENCE 538 AA; 57285 MW; AF28F2EF36413BBD CR-64;

Query March 7 74; Score 164.5; DR 10; Length 538;
Best Local Similarity 24.06; Pred. No. 0.0042;
Matches 99; Conservative 49; Mismatches 146; Indels 125; Gaps 19;

QY 39 SISFLKLIAGKQESVDPDELIDLYGQKKKKDDQTLDFYGTQGSTVYVFKSWPEPD 98
DB : : : : :
DB 39 TVESFKELIA---CNSDVPANQCELYKGRLLKDDCTLLSYGLCACTVHYVGRGVFSSP 95
QY 99 QKP-----EVDKVAAYREP--- --AV 115
DB 96 SAPAANAGNOTTAPQAVGSSNDSNGESLFDGPNLPGGNMAGLFGSLDLECA 155
QY 116 LHTALHSSSYRE---AVFKLSNRES-DQIVATRLSSNPAALGVGLCK-ELRSV 169
DB 156 QQCLAQNPWIREMNTPTAQNLENNPEYRSNKNPQVRE-----LVGNRELGHV 208
QY 169 FADPNM--DTLVA-HPALVNAVLVLHVSAGAPMPGTDSSSPSSSYKDTGPF-- 224
DB 209 LNDSEILQLEAARNPELRM-----MENTD-PAM-SNHSYVEGPNV 251
QY 225 ---LFEGLSDDEDFHNTSTSTSTPSRF--ASGVSGAGRPITCSKATALLALA 279
DB 252 LRMYENVY--CEPLNATTVSEAGNTSNRIALLGNQVITQSDTSNNISAPNAET 309
QY 280 STPESSHTPTQGHSSG--TSVSSGVCSG--- -- --TFITN 315
DB 310 GTPVAN--FLPNPAGATACTAFNENAGLFLNGLNGLNGLNGLNGLNGLNGLN 365
QY 316 DLFSQALQALQASQSSLSLQVQ----- -- -- -- -- -- -- -- -- -- 400
DB 367 QL-SQLLNTFAMSQVSVSNQVWMLMSATQLRQVLEMLNGLREKVVQPF 418

RESULT 13
QY Q9VWD9 PRELIMINARY; PRT, 547 AA;
AC Q9VWD9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created);
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update);
DT 01-MAY-2003 (TrEMBLrel. 21, Last annotation update);
DE CG14224 protein (CG1389:19);
CS CG14224;
GN Drosophila melanogaster (Fruit fly);
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila;
OX NCBI_TaxID=7227;
RN 1;
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

```

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.B.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Bencs P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova J., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferizak C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb C.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kamos I., Simpson V., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector K., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye C., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RC STRAIN=Berkeley;
RA Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez R., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungal C.J., Nunoo C., Pacleb C., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/CDR databases.
DR EMBL: AC003512; AA54903.1;
DR EMBL: AY228447; AA74940.1;
DR HSSP: P02248; 1381;
DR F.Ybase; FBNC031057; CG14224;
DR InterPro: IPR000443; UBA domain.
DR InterPro: IPR006226; Ubiquitin.
DR Pfam: PF00627; UBA; 1.
DR Pfam: PF02240; Ubiquitin; 1.
DR PROSITE: PS00053; Ubiquitin-N2; 1.
SQ SEQUENCE 547 AA; 58836 MW; 8496DC6EC165C4D7 CR-64;

Query March 7 63; Score 161.5; DB 5; Length 547;
Best Local Similarity 22.48; Pred. No. 0.00085;
Matches 100; Conservative 60; Mismatches 170; Indels 117; Gaps 20;

QY 18 TPKSILKLPETELGEYSLOGYSISFTKQLIAGKQESVDPDELIDLYGQKKKKDDQTL 77
DB : : : : :
DB 15 TPKD---KKTVEVDEDS---GKDKFKILVAQKFE---AEPEQLVLFAGKTKMDTTLQ 64
QY 78 FYGIQPGSTVHVLKSKSPEDQKE--EPVD-----KVAAMR----- 111
DB 65 MHNIKDLNLTVELVTKATRNNECPAPQADVROTPTGLNFGGLAGALGAGSNTFMDL 124
QY 112 EFRVLHTALHSSSYRE-----AVFKLSNRESLDQIVATPG----LSSDPTALGVLD 162
DB : : : : :
RX MEDLINE=20196006; PubMed=10731132;

```

DE 125 QARMONELLNNGMLRSYNDPMVQMMNNECTMRQLITSRPMQMDLMQNPESHVLYN 184
 QY 163 KDL-----FSVFADPNMLDTCPAHP-ALVNA:VLVLHVSAG -----SAPM--P 203
 DB 185 PDLRQTVLARNPMQELNRSHPAMN- LSVPGYSALQRIYRDQERPMKA 239
 QY 204 GTDSSSRMPASSYRDMPPGGFLFECGLSDDELCDFHNTSTPSSSPASLQYSGAAG 263
 DB 240 ATESFGRN PRAGIVD--GGSGAGNNPQCTENRNPENFWNANSQTNVTVGGSAKN 296
 QY 264 P-----EPITOSELATALALATPSSSHITPTQGHSSG 299
 DE 297 PTGELPPNVLNTPAIVRSLLQVNAIPAMQRLINARVTRSMVSMQOMAPLUS - 354
 QY 300 TSPMSGVCSGTPITNCLIFSGALQALQASCSRLQS --- --LWQILQQLRD 347
 DB 355 -----SEPLSMNPALQEQVRQXVTFNAGMNPENWMLNFIAMNALILQIQIMELPS 410
 QY 348 M--GIODELSLRCPRPVGTGSKQWS 312
 DB 411 AAPGLV GTLGIPPEPGAGTGNDPAS 436
 RESULT 14
 Q9JUP9 PRELIMINARY; PRT: 582 AA.
 AC Q9JUP9;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created;
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE DA41.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Lung.
 RA Ozaki T., Hishiki T., Nakagawara A., Sakiyama S.
 RT "Identification of a new protein that can associate with DAN."
 RL Submitted (SEP 1996) to the EMBL/GenBank/CCDS databases.
 DR EMBL: D87950; BAA92267.1;
 DR HSSP: P02248; ITB8.
 DR InterPro: IPR006636; STI;
 DR InterPro: IPR000449; UBA domain.
 DE InterPro: IPR000626; Ubiquitin.
 CF Pfam: PF00627; UBA; 1.
 DR Pfam: PF00240; ubiquitin; 1.
 DR SMART: SM00727; STI; 4.
 DR SMART: SM00165; UBA; 1.
 DR SMART: SM00213; UQ;
 DR PROSITE: PSS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 582 AA; 62071 MW; 64558767.1000 CP 1000

Query Match 7.4%; Score 158.5; DB 11; Length 582;
 Best Local Similarity 22.7%; Pred. No. 0.0015;
 Matches 88; Conservative 67; Mismatches 150; Indels 83; Gaps 18;

QY 15 QPTPKSLRPELTELGEYSLGGYSISFLKQLIAGLKQESVDPPELIDIVYGRKLD 73
 DB 24 EPKIMKVTYKPK-EKEEFAPVENSVCQFKELSKFKSQ--DQLVLIFAGK 79
 QY 74 QTLDFYGIQGSTVHLRK-----SAPFPCQKPEFVQVAMNEE 112
 DB 80 DTLSCQGHGLTVHLVIXTKRQDENSACQNTTNSVTSPPAES--NTSGPAANS 137
 QY 113 FRV-----LRTALHSSSSVREAVFM--LKKESLDLIIVATPGLSSDPTALGV 160
 DB 138 FLGGGLGGLAGLSSGLINTTNSFELCSQNRQRLNENPMKLM-----ENFVQSL 190
 QY 161 QDKDLFS--VFADPNMLDTLPAHPALVNA--IVLVHVSAGSAPMPGTSSRSXP 213
 DB 191 SNPDLRQIMANPQW-QQLIQNPFIHVLNPN:MQTLEARNPANYQEMVNRQERD 249

QY 214 SSSYRMPGGF-----LFEGLSDD-----EDDFHNTSTPSSSPASLQYSGAAG 263
 DB 250 LSNLESI-EGYNALRMYTDIQEPLNAAQEQFGN-----PFASLVSP-----SSNRG 299
 QY 264 PRPITOSELATALALATPSSSHITPTQGHSSGTPSMSSGVOSGTPITNCLIFSOALQ 323
 DB 300 TQP-SRTENRDPNPAPQTPQSSPASGSGTGTINTVTSAGNATSTPAQCGTSGPNLV 358
 QY 324 HALQAS--GQPSLOSQWQ-----PCLQQ 344
 DB 359 PGAGASMTPTGMSLQOITENPOLYC 386
 RESULT 15
 Q9LHD9 PRELIMINARY; PRT: 624 AA.
 AC Q9LHD9;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created;
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ubiquitin-like product Chap./Dsk2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=20141258; PubMed=10675567;
 RA Kaye F.J., Modi S., Ivanovska I., Koonin E.V., Thress K., Kubo A.,
 RA Kornbluth S., Rose M.D.;
 RT "A family of ubiquitin-like proteins binds the Affase domain of Hep70-
 RT like Stch."
 RL FEBS Lett. 467:348-355(2000).
 DR EMBL: AF189509; AAF17237.1;
 DR HSSP: P02248; ITB8.
 DR InterPro: IPR006636; STI;
 DR InterPro: IPR000449; UBA domain.
 DE InterPro: IPR000626; Ubiquitin.
 CF Pfam: PF00627; UBA; 1.
 DR Pfam: PF00240; ubiquitin; 1.
 DR SMART: SM00727; STI; 4.
 DR SMART: SM00165; UBA; 1.
 DR SMART: SM00213; UQ; 1.
 DR PROSITE: PSS0053; UBIQUITIN 2; 1.
 SQ SEQUENCE 624 AA; 65765 MW; DEAT53B50DC6AC3 CRC64;

Query Match 7.1%; Score 152.5; DB 4; Length 624;
 Best Local Similarity 19.8%; Pred. No. 0.0048;
 Matches 97; Conservative 70; Mismatches 177; Indels 147; Gaps 19;

QY 13 AQPLTPKSLRPELTELGEYSLGGYSISFLKQLIAGLKQESVDPPELIDIVYGR 68
 DB 24 AAAPAEK:IKVTYKPK-EKEEFAPVENSVCQFKELSKFKSQ--DQLVLIFAGK 79
 QY 69 KLKDDQTLDFYGIQGSTVHLRKSHPEPD-QKPEVDKVAAMREFVHLTHALHSSSYR 127
 DB 80 ILKDDQTLDFYGIQGSTVHLRKSHPEPD-QKPEVDKVAAMREFVHLTHALHSSSYR 130
 QY 128 EAVFMKLSN-----KESLDLIIVATP-----GLSSDPI 155
 DB 131 SNSTP-STNSNPFGLSLGGLAGLSSGLSSTNFSLOSOQOOLMASPEMIQIMENPF 190
 QY 156 ALGVQCKDLFS--VFADPNMLDTLPAHPALVNA-----IVLVHVSAGSAPMPGTSS 208
 DB 191 VQSMLSNPDLRQIMANPQW-QQLIQNPFIHVLNPN:MQTLEARNPANYQEMVNR 249
 QY 209 SRSMPSSSYRDMPPGGF-LFEGLSDDEDFHNTSTPSSSPASLQYSGAAG----- 263
 DB 250 NQDLALSNLESTPGGYNALRRMTDIOEPLNAAQEQFGNPF-----ASVSSSSSGGTQ 306
 QY 264 -----PRPITOSELATALALATPSSSHIT-----P 289

Db	307	PSRTENDPLNPKAPPATQSS-AITSTTTSTGSGNSSNATONTVAANYVASIFS	365
Qy	290	TPGTQG-----HSSGTSPMSSGVQSGTPTNDLFSQALQHA-QASGQPSLQSQ	337
Db	366	TPGVCSILQQITENPQLIQKXLSAPYKSMWQSLSQNFDLAAQYMLNSPLFTANPQJQEQ	425
Qy	338	WQPOLQQLRDMGQDDELSURPQPPVGTGSKQFWSSSLIEEHEHLPASPEPPASGQYCP	397
Db	426	MRPQL-----EAFQCMQNTGUSAYNPPA	451
Qy	398	WEALYKYVPSL	408
Db	452	MQALYQIQQGL	462

Search completed: November 12, 2003, 04:34:34
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 06:42:23 : Search time 46 seconds
(without alignment)
14:41:37 Million cell updates/sec

Title: US-09-735-251-4

Perfect score: 2133

Sequence: 1 MSLSHHLAVKLADQLTPK ... SGPVCEWALMVETLVY 410

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 15972453 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03

1	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
2	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
3	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
4	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
5	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
6	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
7	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
8	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
9	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
10	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
11	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
12	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
13	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
14	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
15	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
16	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
17	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
18	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
19	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
20	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
21	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
22	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
23	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT
24	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the best being found, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2133	100.0	410	19	Activated platelet
2	2129	99.8	426	18	Human activated pi
3	1834	86.0	420	20	Amino acid sequenc
4	1834	86.0	420	22	Human secreted pro
5	1834	86.0	420	23	Human polypeptide
6	1828	85.7	380	22	Human protein HPC
7	1686	79.0	341	23	Human polypeptide
8	794	37.2	191	22	Human prostatic can
9	592	27.8	164	22	Human prostate can

10	566	26.5	128	22	AA013489	Human polypeptide
11	389	18.2	76	21	AAG02759	Human secreted pro
12	370	17.3	114	21	AAB58780	Breast and ovarian
13	164.5	7.7	536	21	AAG30040	Arabidopsis thalia
14	164.5	7.6	577	21	AAG30039	Arabidopsis thalia
15	161.5	7.6	547	22	ABG9125	Drosophila melanog
16	152.5	7.1	624	22	AAB47122	Human Chapi. Homo
17	152.5	7.1	624	23	ABG97173	Novel human protei
18	152	7.1	329	23	ABJ10042	Yeast selected int
19	149	7.0	415	22	ABG59727	Drosophila melanog
20	145.5	6.8	589	22	AAJ94008	Human stomach canc
21	145.5	6.8	589	22	AAB95205	Human protein sequ
22	141.5	6.6	589	22	AAB94311	Human protein sequ
23	141.5	6.6	589	22	AAB62194	Human SAP2 amino a
24	141.5	6.6	604	22	ABG16557	Novel human diagno
25	140.5	6.6	589	22	ABU32717	Human brain-derive
26	139.5	6.5	1093	16	ABG66460	AF-17 protein. Ho
27	139.5	6.5	1093	16	ABG66460	Human translocatio
28	139	6.5	589	22	AAB74670	Human protease and
29	136.5	6.4	464	21	AAB72447	Human 149Y2H151 u
30	134	6.3	655	21	AAJ95624	Human Zalpha16 pro
31	134	6.3	655	24	ABG73191	Human Zalpha16 pro
32	134	6.3	655	24	ABJ19251	Human cancer/testi
33	134	6.3	1145	22	AAU04895	Microspora eve
34	134	6.3	1963	22	ABG2919	Drosophila melanog
35	133.5	6.3	902	24	ABU4281	Human expressed pr
36	133.5	6.3	902	24	ABU4281	Human expressed pr
37	133.5	6.3	907	24	ABJ04278	Human expressed pr
38	133.5	6.3	907	24	ABU04306	Human expressed pr
39	133.5	6.3	907	24	ABU04308	Human expressed pr
40	133.5	6.3	2311	23	ABG5034	Human translocatio
41	133	6.2	1012	20	AAJ17406	Human atrophin-1 r
42	133	6.2	1168	22	ABG02040	Novel human diagno
43	133	6.2	1487	23	AAO14356	Equine herpesvirus
44	132	6.2	4899	22	ABG65895	Drosophila melanog
45	131.5	6.2	773	22	ABG15668	Novel human diagno

ALIGNMENTS

RESULT 1
AAW15414
ID AAW15414 standard; Protein: 410 AA.
XX AC AAW15414;
XX DT 06-JUN-1997 (first entry)

DE Activated platelet protein-2 APP-2 alternatively spliced variant.
XX APP-2; activated platelet protein-2; thrombus; monoclonal antibody.
XX Homo sapiens.

XX WO9712898-A1.
XX 10-APR-1997.
XX 04-OCT-1996; 96WO-US15922.
XX 06-OCT-1995; 95US-0005074.
XX (HARD) HARVARD COLLEGE.
XX Clement CY, Reed G;
XX WPI; 1997-226153/20.
XX N-PSDB; AAT64550.

XX Naturally-occurring platelet activation polypeptide, APP-2 - used
PT for locating a platelet thrombus in an animal, or activated platelet
PT complex in a biological sample

XX Claim 2; Fig 6; 50pp; English.

PS A polypeptide (AAW15413) comprises an alternatively spliced variant

XX of novel human activated platelet protein-2 (APP-2) (see also

CC AAW15413), a protein that is preferentially expressed on activated

CC human platelets but not resting platelets. Its amino acid sequence

CC was deduced from an isolated cDNA clone (AAW64550). APP-2 is

CC characterised as confg. an epitope that binds to monoclonal antibody

CC 3B2 (ATCC CRL 11986). APP-2 can be produced in transformed host

CC cells. Anti-APP-2 antibodies are useful in methods for detecting

CC activated platelets in biological samples, for localising a platelet

CC thrombus in an animal, and for detecting an activated platelet

CC protein complex in an animal.

XX SQ Sequence 410 AA;

Query Match 100.0%; Score 2132; DB 18; Length 410;

Best Local Similarity 100.0%; Pred. No. 5.1e-167;

Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSDHHLAVKLAADQPLTPKSIILRLPETELGEVSLGGYSISFLKQIAGKQESVDPPEL 60

DB 1 MSLSDHHLAVKLAADQPLTPKSIILRLPETELGEVSLGGYSISFLKQIAGKQESVDPPEL 60

QY 61 IDLVCGRKLKDDQTLDFYGIQPGSTVHVLRKSNPEPDQKPEVDKYAAMREFRVLHTAL 120

DB 61 IDLVCGRKLKDDQTLDFYGIQPGSTVHVLRKSNPEPDQKPEVDKYAAMREFRVLHTAL 120

QY 121 HSSSSYREAVFVLMKSNKESLDQIIVATPGLSSDPDIALGVLDKLFVSFAQPNMLDTLVP 180

DB 121 HSSSSYREAVFVLMKSNKESLDQIIVATPGLSSDPDIALGVLDKLFVSFAQPNMLDTLVP 180

QY 181 AHPALVNAIVLVLSVAGSAPMPGTGSSSRMPSSSYRDMPGFELFEGLSDDDEDFHPT 240

DB 181 AHPALVNAIVLVLSVAGSAPMPGTGSSSRMPSSSYRDMPGFELFEGLSDDDEDFHPT 240

QY 241 RSTPSSSTFSSRPASLGYSAGAPRPITQSELATALALASTPSSSHTPTPGTQGHSSGT 300

DB 241 RSTPSSSTFSSRPASLGYSAGAPRPITQSELATALALASTPSSSHTPTPGTQGHSSGT 300

QY 301 SPSSGVSQSGTPIITNLFQQA-CHALQALQASQPSLSQSQWQFQQLQRMGICQDDELSRPC 360

DB 301 SPSSGVSQSGTPIITNLFQQA-CHALQALQASQPSLSQSQWQFQQLQRMGICQDDELSRPC 360

QY 361 RPPVGTSGKQWSSSLLEPHELPAPEPPASCRGCPWEALMKVPPSLP 409

DB 361 RPPVGTSGKQWSSSLLEPHELPAPEPPASCRGCPWEALMKVPPSLP 409

RESULT 2

AAW15413

IJ AAW15413 standard; Protein; 426 AA.

XX AAW15413;

CC 26-JUL-1997 (first entry)

XX Human activated platelet protein 2 APP 2.

DE APP 2; activated platelet protein-2; thrombus; monoclonal antibody

XX Homo sapiens.

CS W09712898-A1.

EN 10-APR-1997.

XX 04-OCT-1996; 96WO-US15922.

XX 06-OCT-1995; 95US-0005074.

XX (HARC) HARVARD COLLEGE.

XX Clement CY, Reed G;

XX WPI; 1997-226153/20.

DR N-PSDB; AAW164549

XX Naturally-occurring platelet activation polypeptide, APP-2 - used

PT for locating a platelet thrombus in an animal, or activated platelet

PT complex in a biological sample

XX Claim 1; Fig 5; 50pp; English.

PS A novel 25 kDa polypeptide (AAW15413) designated activated platelet

CC protein-2 (APP-2) is preferentially expressed on activated human

CC platelets but not resting platelets. APP-2 is characterised as

CC confg. an epitope that binds to monoclonal antibody 3B2 (ATCC CRL

CC 11986). Its amino acid sequence was deduced from a cDNA clone

CC (AAW64549) obtd. from a human bone marrow cDNA library. An

CC alternatively spliced variant (see also AAW15414) was also isolated.

CC APP-2 can be produced in transformed host cells. Anti-APP-2

CC antibodies are useful in methods for detecting activated platelets

CC in biological samples, for localising a platelet thrombus in an

CC animal, and for detecting an activated platelet protein complex in

CC an animal.

XX SQ Sequence 426 AA;

Query Match 99.9%; Score 2129; DB 18; Length 426;

Best Local Similarity 100.0%; Pred. No. 5.1e-167;

Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSDHHLAVKLAADQPLTPKSIILRLPETELGEVSLGGYSISFLKQIAGKQESVDPPEL 60

DB 1 MSLSDHHLAVKLAADQPLTPKSIILRLPETELGEVSLGGYSISFLKQIAGKQESVDPPEL 60

QY 61 IDLVCGRKLKDDQTLDFYGIQPGSTVHVLRKSNPEPDQKPEVDKYAAMREFRVLHTAL 120

DB 61 IDLVCGRKLKDDQTLDFYGIQPGSTVHVLRKSNPEPDQKPEVDKYAAMREFRVLHTAL 120

QY 121 HSSSSYREAVFVLMKSNKESLDQIIVATPGLSSDPDIALGVLDKLFVSFAQPNMLDTLVP 180

DB 121 HSSSSYREAVFVLMKSNKESLDQIIVATPGLSSDPDIALGVLDKLFVSFAQPNMLDTLVP 180

QY 181 AHPALVNAIVLVLSVAGSAPMPGTGSSSRMPSSSYRDMPGFELFEGLSDDDEDFHPT 240

DB 181 AHPALVNAIVLVLSVAGSAPMPGTGSSSRMPSSSYRDMPGFELFEGLSDDDEDFHPT 240

QY 241 RSTPSSSTFSSRPASLGYSAGAPRPITQSELATALALASTPSSSHTPTPGTQGHSSGT 300

DB 241 RSTPSSSTFSSRPASLGYSAGAPRPITQSELATALALASTPSSSHTPTPGTQGHSSGT 300

QY 301 SPSSGVSQSGTPIITNLFQQA-CHALQALQASQPSLSQSQWQFQQLQRMGICQDDELSRPC 360

DB 301 SPSSGVSQSGTPIITNLFQQA-CHALQALQASQPSLSQSQWQFQQLQRMGICQDDELSRPC 360

QY 361 RPPVGTSGKQWSSSLLEPHELPAPEPPASCRGCPWEALMKVPPSLP 409

DB 361 RPPVGTSGKQWSSSLLEPHELPAPEPPASCRGCPWEALMKVPPSLP 409

RESULT 3

AAW42382

ID AAW42382 standard; Protein; 420 AA.

XX AAW42382;

CC 09-DEC-1999 (first entry)

XX Amino acid sequence of fx317_11.

DE secreted protein; cDNA library; clone; transmembrane protein;

XX signal sequence cloning; hybridization cloning; gene therapy;

KW receptor.

QY	241	RSTSSSTPSSRPASLGYSGAAGPRP;TQSELATATALASTPSSSHPTTGTQGHSSST	300
DB	281	RSTSSSTPSSRPASLGYSGAAGPRP;TQSELATATALASTPSSSHPTTGTQGHSSST	340
QY	301	SPMSGVQSGTPTINDLFSQALQHALQASQPSLOSQWQPOLQQLRDMGIQDELSLR	358
DB	341	SPMSGVQSGTPTINDLFSQALQHALQASQPSLOSQWQPOLQQLRDMGIQDELSLR	398
RESULT 4			
AAU39074		ID	AAU39074 standard; Protein: 420 AA.
XX	AC	AAU39074;	
XX	DT	16-JAN-2002 (first entry)	
XX	XX	Human secreted protein fx317_11.	
XX	XX	Human; secreted protein; anti-inflammatory; immunosuppressive; neurotropic; neuroprotective; antiarthritic; antimicrobial; vulnery; cystostatic; antidiabetic; virucide; antiinfertility; anticonvulsant; vasotropic; antiparkinsonian; immunostimulant; dermatological; antirheumatic; antitumor; antitumor; osteopathic; tranquiliser; cerebroprotective; cytokine; cell proliferation; cell differentiation; immune deficiency; severe combined immunodeficiency; SCID; tumour; autoimmune disorder; multiple sclerosis; rheumatoid arthritis; graft-versus-host disease; myeloid deficiency; wound healing; ulcer; periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease; Parkinson's disease; Huntington's disease; infection; cardiac disease; stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen; food supplement; vaccine.	
XX	OS	Homo sapiens.	
XX	PN	WO200175068-A2.	
XX	PD	11-OCT-2001.	
XX	XX	22-MAR-2001; 2001WO-US09369.	
XX	PR	30-MAR-2000; 2000US 0539330.	
XX	PR	04-DEC-2000; 2000US-0729674.	
XX	PA	(GENY) GENETICS INST INC.	
XX	PI	Jacobs K, McCoy JM, Lavallie E, Collins-racie LA, Evans C, Treacy M, Acostino MJ, Steininger RJ, Spaulding V, Wong GG; Clark H, Pechter K, Nerberg D;	
XX	DR	WPI: 2001-629363/73.	
XX	DR	N-PSDB: AAS59292.	
XX	PT	Secreted human proteins, useful as vaccine for treating various diseases such as autoimmune disorders (e.g. multiple sclerosis), and nervous system disorders (e.g. stroke) -	
XX	XX	Disclosure: Page 578-579: 619pp; English.	
XX	XX	The invention relates to novel human secreted proteins, the nucleic acids encoding them. The protein may exhibit cytokine, cell proliferation or cell differentiation activity or may induce production of other cytokines in certain cell populations and may exhibit immune stimulating or immune suppressing activity, which is useful for the treatment of various immune deficiencies and disorders e.g. severe combined immunodeficiency (SCID), autoimmune disorders e.g. multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation. The proteins are also useful in the treatment of diseases and disorders including tissue, skin and organ transplantation and in graft-versus-host diseases (GVHD), in the induction of tumour immunity, myeloid or lymphoid cell deficiencies, wound healing and tissue repair, in the treatment of burns, incisions and ulcers; as well as in treatment of periodontal disease, osteoporosis or osteoarthritis, mediated by	

CC	inflammatory processes, diseases of the peripheral nervous system.
CC	Alzheimer's, Parkinson's disease, Huntington's disease.
CC	amyotrophic lateral sclerosis, and Shy-Drager syndrome, infections,
CC	infection of cardiac and central nervous system vessels, e.g. stroke,
CC	sepsis, inflammatory bowel disease, ulcers, bone regeneration, the
CC	protein, having activin- or inhibin-related activities is useful as a
CC	contraceptive based on the ability of inhibins to decrease fertility in
CC	female mammals and decrease spermatogenesis in male mammals. The
CC	proteins and nucleic acids are also useful as food supplements. The
CC	present sequence represents a secreted protein of the invention
XX	
SQ	Sequence 420 AA:
Query Match	
Best Local Similarity 100.0%; Score 1834; DB 23; Length 420;	
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSLSDMHLAVKLADQPLTPKSIILRLPETELGYSYSGYSLFSLKOLIAKQESVPDEP 60
DB	4: MSLSDMHLAVKLADQPLTPKSIILRLPETELGYSYSGYSLFSLKOLIAKQESVPDEP 100
QY	61 IDLIYGRKLLKDDQTLFYSYIQGHTVHLVLRWPEQPKPEYKVKVAKRPPVILTA 120
DB	101 IDLIYGRKLLKDDQTLFYSYIQGHTVHLVLRWPEQPKPEYKVKVAKRPPVILTA 169
QY	121 HSSSYREAVFMKLSNKESLQIVATPGLSSPITALGVLCQKLFYSVFAPKPYLDTAVP 180
DB	161 HSSSYREAVFMKLSNKESLQIVATPGLSSPITALGVLCQKLFYSVFAPKPYLDTAVP 220
QY	181 AHPALVVAIVLHVSAGSAPMPTLSSRSVSSSYRMPDPEKPEGLSGDEDFEPT 240
DB	221 AHPALVVAIVLHVSAGSAPMPTLSSRSVSSSYRMPDPEKPEGLSGDEDFEPT 280
QY	241 RSTPSSSTPSRPSASLOYSAGAGPRTTQSEATATLALASTPESSSHTPTGTQGHSSGT 300
DB	281 RSTPSSSTPSRPSASLOYSAGAGPRTTQSEATATLALASTPESSSHTPTGTQGHSSGT 340
QY	301 SPMSGVSQSGTPTINDLFSQALCHALCASQSLSSQWQPCQGLQREXGICFHLSP 358
DB	341 SPMSGVSQSGTPTINDLFSQALCHALCASQSLSSQWQPCQGLQREXGICFHLSP 398
RESULT 5	
AS555783	
ID	AB855783 standard; Protein; 420 AA
AC	
XX	AB855783;
XX	
DT	14-FEB-2002 first entry.
XX	
DE	Human: polypeptide SEQ ID NO 1/2
XX	
KW	Human; clone bd306-7; clone YB8 is ATCC number 9859; gene therapy.
KW	immune disorder; bacterial infections; fungal infections; cancer; tumour;
KW	autoimmune disorder; systemic lupus erythematosus; wound; ulcer; inhibin;
KW	osteoporosis; osteoarthritis; nervous system disorder; neuropathy;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; activin;
KW	haemophilia; cardiac infarction; stroke; sepsis; arthritis; vulvectomy;
KW	ischaemia-reperfusion injury; inflammatory bowel disease; chemotactic;
KW	Crohn's disease; cytostatic; anti-inflammatory; immunomodulatory;
KW	neuroprotective; haemostatic; thrombolytic; anti-inflammatory.
XX	
CS	Homo sapiens.
XX	
PN	US2001039335-A1.
XX	
PD	
XX	
XX	09-NOV-2003.
XX	
FF	04-DEC-2000; 2000US-0729674.
XX	
PR	10-APR-1997; 97US-126425P.
PR	04-DEC-1997; 97US-067454P.
IR	20-DEC-1997; 97US-068179P.

PR	02-JAN-1998; 98US-070346P.
PR	07-JAN-1998; 98US-070643P.
PR	08-JAN-1998; 98US-070755P.
PR	13-JAN-1998; 98US-071304P.
PR	22-JAN-1998; 98US-072134P.
PR	30-JAN-1998; 98US-073095P.
PR	18-FEB-1998; 98US-075038P.
PR	30-MAR-2000; 2000US-053933C.
PR	23-NOV-1998; 98US-019785P.
XX	
PA	(CACO/) JACOBS K.
PA	(MCCO/) MCCOY J M.
PA	(LAVA/) LAVALLIER E W.
PA	(COLL/) COLLINS RALIE L A.
PA	(EVAN/) EVANS C.
PA	(MERB/) YERBERG D.
PA	(TREA/) TREACY M.
PA	(AGOS/) AGOSTINO V J.
PA	(STEI/) STEININGER P J.
PA	(SPAU/) SPAULDING V.
PA	(WONG/) WONG G G.
PA	(CLAR/) CLARK H.
PA	(FECH/) FECHTEL K.
XX	
PI	Jacobs K, McCoy JM, Lavallier ER, Collins-Racie LA, Evans C,
PI	Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;
PI	Wong GG, Clark H, Fechtel K;
XX	
DR	WPI: 2002-040725/05.
DR	N-PSDB; ABA90962.
XX	
PT	New secreted proteins and encoding polynucleotides, useful in gene
PT	therapies, particularly for preventing or treating autoimmune
PT	disorders, cancer, graft-versus-host disease, wound, osteoporosis,
PT	stroke or inflammations
XX	
PS	Disclosure; Page 309-310; 349pp; English.
XX	
CC	The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
CC	ABA90980) and encoded proteins (AB855699-AB855800), especially
CC	polynucleotides SEQ ID NO 1 (ABA90876) and SEQ ID NO 19 (ABA90885) and
CC	proteins SEQ ID NO 2 (AB855698) and SEQ ID NO 20 (AB855707) contained in
CC	clones bd306-7 and YB8-1 respectively and the clones bd306-7 and YB8-1
CC	are deposited with the American Type Culture Collection (ATCC) with
CC	accession number 98599. The polynucleotides and encoded polypeptides have
CC	cytostatic, anti-inflammatory, immunomodulator, vulvectomy, thrombolytic
CC	and neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
CC	and anti-inflammatory activity and acting as cytokine modulators,
CC	haematopoietic regulators, tissue growth modulators and/or cadherin
CC	suppressors. The polypeptides and polynucleotides are useful in gene
CC	therapies, particularly for preventing, treating or ameliorating any of
CC	the following diseases: immune deficiency and disorders; e.g. bacterial
CC	or fungal infections, autoimmune disorders, cancer, systemic lupus
CC	erythematosus or graft-versus-host disease; myeloid or lymphoid cell
CC	deficiencies; wound, burns, incisions and ulcers, osteoporosis or
CC	osteoarthritis; central and peripheral nervous system diseases and
CC	neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
CC	disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
CC	haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
CC	or systemic inflammatory response syndrome, ischaemia-reperfusion
CC	injury, endotoxin lethality, arthritis, inflammatory bowel disease or
CC	Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
CC	foliaceus.
XX	
SQ	Sequence 420 AA:
Query Match	
Best Local Similarity 100.0%; Score 1834; DB 23; Length 420;	
Matches 358; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MSLSDMHLAVKLADQPLTPKSIILRLPETELGYSYSGYSLFSLKOLIAKQESVPDEP 60
DB	4: MSLSDMHLAVKLADQPLTPKSIILRLPETELGYSYSGYSLFSLKOLIAKQESVPDEP 100

QY 61 IDLIYCGKGLKDDQTLDFYGTQPGSTVHVLKSNPEPCQKPEPVCKVAAMREFRVLTAL 120
 DB 101 IDLIYCGKGLKDDQTLDFYGTQPGSTVHVLKSNPEPCQKPEPVCKVAAMREFRVLTAL 160
 QY 121 HSSSYRAVFKMLSNKESLDQIIIVATPGLSQALGVACQKDLFSVFALPNMLDTLVP 180
 DB 161 HSSSYRAVFKMLSNKESLDQIIIVATPGLSQALGVACQKDLFSVFALPNMLDTLVP 220
 QY 181 AHPALVNAIVLVLSVAGSAMPPTDSSSRSPSSSYRDMFGFLFEGLSDDDDFHNT 240
 DB 221 AHPALVNAIVLVLSVAGSAMPPTDSSSRSPSSSYRDMFGFLFEGLSDDDDFHNT 280
 QY 241 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTPGTCHSSGT 300
 DB 281 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTPGTCHSSGT 340
 QY 301 SPMSSGVOSGTPITNCLFSQALCALQASGQPSQSQWQPOLQQLRDMGICDDLSLR 358
 DB 341 SPMSSGVOSGTPITNCLFSQALCALQASGQPSQSQWQPOLQQLRDMGICDDLSLR 396

RESULT 6
 AAG93292
 ID AAG93292 standard; Protein; 380 AA.
 AC AAG93292;
 DT 13-SEP-2001 (first entry;
 XX Human protein: HP10657.
 DE Human; gene therapy; tumour.
 KW Homo sapiens.
 OS Homo sapiens.
 XX WO200142302-A1.
 PC 14-JUN-2001.
 FE 06-DEC-2000; 2000ACJP08631.
 XX 06-DEC-1999; 99JP-0346663.
 PR 06-DEC-1999; 99JP-0346664.
 PS 08-FEB-2000; 2000JP-0031062.
 PS 10-FEB-2000; 2000JP-0034090.
 PS 10-FEB-2000; 2000JP-0034091.
 PS 14-FEB-2000; 2000JP-0035829.
 PR 14-FEB-2000; 2000JP-0035849.
 PR 14-MAR-2000; 2000JP-0035849.
 PR 30-MAY-2000; 2000JP-0035849.
 XX (INIS) JAPAN SCI & TECHNOLOGY DATA
 PA Kato S, Eguchi C, Saeki M,
 PI WPI; 2001-381646/40.
 DR N-PSDB; AAH6R577.
 XX Human protein originated from tumour cell line, applicable as drug,
 PT reagent for studying intracellular protein networks and protein source
 PT for drug screening, also encoded data for gene diagnosis and gene
 PT therapy.
 XX Claim 1; Pages 313-314; 47pp; Japanese.
 PS The present sequence is a human protein, the human protein, preferably
 CC originated from tumour cell line, is applicable as a drug, a reagent for
 CC studying intracellular protein networks and a protein source for
 CC screening proteins for binding low molecular weight drugs, the human
 CC protein coding sequence is useful for gene diagnosis and gene therapy,
 CC expression vectors and transformant cells for detection of ligands and
 CC receptors.

XX SQ Sequence 380 AA;
 Query Match 85.7%; Score 1828; DB 22; Length 380;
 Best Local Similarity 99.7%; Pred. No. 2.9e-142;
 Matches 357; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MSLSNWHAVKLAQCPITPKSLIRLPETELCEYSLGGYSISFLKQLIAGKLCRSVPDPPEL 60
 DB 1 MSLSNWHAVKLAQCPITPKSLIRLPETELCEYSLGGYSISFLKQLIAGKLCRSVPDPPEL 60
 QY 61 IDLIYCGKGLKDDQTLDFYGTQPGSTVHVLKSNPEPCQKPEPVCKVAAMREFRVLTAL 120
 DB 61 IDLIYCGKGLKDDQTLDFYGTQPGSTVHVLKSNPEPCQKPEPVCKVAAMREFRVLTAL 120
 QY 121 HSSSYRAVFKMLSNKESLDQIIIVATPGLSQALGVACQKDLFSVFALPNMLDTLVP 180
 DB 121 HSSSYRAVFKMLSNKESLDQIIIVATPGLSQALGVACQKDLFSVFALPNMLDTLVP 180
 QY 181 AHPALVNAIVLVLSVAGSAMPPTDSSSRSPSSSYRDMFGFLFEGLSDDDDFHNT 240
 DB 181 AHPALVNAIVLVLSVAGSAMPPTDSSSRSPSSSYRDMFGFLFEGLSDDDDFHNT 240
 QY 241 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTPGTCHSSGT 300
 DB 241 RSTPSSSTPSSRPASGLYSGAAGRPITQSELATALASTPSSSHTPTPGTCHSSGT 300
 QY 301 SPMSSGVOSGTPITNCLFSQALCALQASGQPSQSQWQPOLQQLRDMGICDDLSLR 358
 DB 301 SPMSSGVOSGTPITNCLFSQALCALQASGQPSQSQWQPOLQQLRDMGICDDLSLR 358

RESULT 7
 ABB89944
 ID ABB89944 standard; Protein; 341 AA.
 AC ABB89944;
 DT 24-MAY-2002 (first entry;
 XX Human polypeptide SEQ ID NO 2320.
 DE Cystostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 OS Homo sapiens.
 XX WO200190304-A1.
 FN 29-NOV-2001.
 XX 19-MAY-2001; 2001WC-US16440.
 XX 19-MAY-2000; 2000US-205515P.
 PR (HUMAN) HUMAN GEMOVE SCI INC.
 PA Birse CE, Rosen CA;
 PI WPI; 2002-122026/16.
 DR N-PSDB; ABL9C353.
 XX Novel; 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX Claim 11; SEQ ID NO 2320; 2081pp + Sequence Listing; English.
 PS The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC

(ABB89040-ABB9044) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and anti-agonists are useful in the diagnosis, treatment and prevention of (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. autoimmune disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral amyloid and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in an alternative form directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 341 AA:

Query Match 79.0%; Score 1686; DB 23; Length 341;
 Best Local Similarity 98.2%; Pred. No. 1.1e-135;
 Matches 330; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NSLSQMLAVKLADQFLPKSLRPELDELGYSLGYSISFLKLIAGKIQESVPOFEL 60
 DB : NSLSQMLAVKLADQFLPKSLRPELDELGYSLGYSISFLKLIAGKIQESVPOFEL 60
 QY 61 IDLVYCGKAKDQDTLPYGIQITVHVLRPSNPPEQKTEYDKVAAYRFRVLRTAL 120
 DB IDLVYCGKAKDQDTLPYGIQITVHVLRPSNPPEQKTEYDKVAAYRFRVLRTAL 120
 QY 121 HSSSYRAVAFVYLNKESLDQIVATGSLNPAALGVQKQLFSYFAELNMLDTLP 180
 DB HSSSYRAVAFVYLNKESLDQIVATGSLNPAALGVQKQLFSYFAELNMLDTLP 180
 QY 181 ALPALVNAIVLVHVSAGAPMEKDSRRVPSRSSYEDMEYTFDFEGLMDEGHNT 240
 DB ALPALVNAIVLVHVSAGAPMEKDSRRVPSRSSYEDMEYTFDFEGLMDEGHNT 240
 QY 241 RSTPSSSTPSSRPASIGVSGAARFETQSEKATALASTSSSTPTTQDSHST 300
 DB RSTPSSSTPSSRPASIGVSGAARFETQSEKATALASTSSSTPTTQDSHST 300
 QY 301 SPYSSVQSGTPTINDLESQALCHLQASLQESL 316
 DB SPYSSVQSGTPTINDLESQALCHLQASLQESL 316

RESULT 9
 AAB63948
 ID AAB63948 standard; Protein: .91 AA
 XX AAB63948;
 AC AAB63948;
 XX 26-MAR-2001 (first entry)
 DT Human prostate cancer associated antigen protein sequence SEQ ID NO:1310.
 DE Human prostate cancer associated antigen; prostate cancer; diagnosis;
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WC2000073801-A2.
 PN 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US14749.
 XX 28-MAY-1999; 99US-0136526.
 XX 10-SEP-1999; 99US-0153454.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX

(LUDW-) LUDWIG INST CANCER RES.
 XX Obata Y;
 XX WPI; 2001-025274/21.
 DR Nucleic acids encoding breast, gastric and prostate cancer associated
 PT antigen precursors, useful for diagnosing and treating a condition
 PT characterized by expression of an abnormal amount of a protein, e.g.
 PT cancer
 XX Example 1; Page 1-11, 799pp; English.
 XX AAF22422 to AAF22426, AAF22627 to AAF22773 and AAF22774 to AAF23014
 CC represent nucleic acid sequences encoding human breast, gastric and
 CC prostate cancer associated antigen precursors (CAAP) respectively.
 CC AAB63232 to AAB63447, AAB63468 to AAB63721 and AAB63722 to AAB63970
 CC represent human breast, gastric and prostate CAAP protein sequence
 CC respectively. CAAPs have cytostatic activity and can be used in the
 CC production of cancer vaccines. The human CAAP proteins, peptides, nucleic
 CC acids or anti-CAAP antibodies are useful for diagnosing and treating a
 CC condition characterized by expression of an abnormal amount of a protein,
 CC e.g. cancer.
 XX SQ Sequence 191 AA:

Query Match 37.2%; Score 794; DB 22; Length 191;
 Best Local Similarity 84.2%; Pred. No. 2.1e-57;
 Matches 165; Conservative 2; Mismatches 15; Indels 14; Gaps 2;

QY 96 EPQCKPEPVQKVAAMREFRVJHTALHSSSSSYREAVFKMLSNKESLDQIVATPGLSSDPI 155
 DB 1 EPQCKPEPVQKVAAMREFRVJHTALHSSSSSYREAVFKMLSNKESLDQIVATPGLSSDPI 60
 QY 156 ALGVLOKQLFSYFAELNMLDTLPAPALVNAIVLVHVSAGAPMPCTDSSSRMPSS 215
 DB ALGVLOKQLFSYFAELNMLDTLPAPALVNAIVLVHVSAGAPMPCTDSSSRMPSS 120
 QY 216 SYRCMGGQLFEGLSDDEDFHNTSTPSSSTPSSRPASLQYSGAAGRPPTQSELATA 275
 DB SYRCMGGQLFEGLSDDEDFHNTSTPSSSTPSSRPASLQYSGAAGRPPTQSELATA 275
 QY 276 LALASTPSSSSTPTPT 291
 DB LALASTPSSSSTPTPT 182

RESULT 9
 AAB63949
 ID AAB63949 standard; Protein: 164 AA
 XX AAB63949;
 AC AAB63949;
 XX 26-MAR-2001 (first entry)
 DT Human prostate cancer associated antigen protein sequence SEQ ID NO:1311.
 DE Human prostate cancer associated antigen; prostate cancer; diagnosis;
 KW Human; breast cancer; gastric cancer; prostate cancer; diagnosis;
 KW cancer associated antigen; cytostatic; cancer vaccine.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WC2000073801-A2.
 PN 07-DEC-2000.
 XX 26-MAY-2000; 2000WO-US14749.
 XX 28-MAY-1999; 99US-0136526.
 XX 10-SEP-1999; 99US-0153454.
 XX (LUDW-) LUDWIG INST CANCER RES.
 XX

[illegible]

PI Data Y:
XX WPI; 2001-025274/03.
DR Nucleic acids encoding breast, gastric and prostate cancer associated
XX antigen precursors, useful for diagnosing and treating a condition
PT characterized by expression of an abnormal amount of a protein, e.g.
Pt Cancer -
PS Example 1; Page 789; 79pp; English.
PPS

CC AAF22422 to AAF22626, AAFA22627 to AAFA22773 and AAFA22774 to AAFA23014
represent nucleotide sequences encoding human breast, gastric and
prostate cancer associated antigen precursors (AAFA21 respectively,
AAB63232 to AAS63467, AAS63468 to AAG63721 and AAG63722 to AAG63790
respectively). CAAPs have cytostatic activity and can be used in the
production of cancer vaccines. The human CAAP proteins, peptides, nucleic
acids or anti-CAAP antibodies are useful for diagnosing and treating a
condition characterised by expression of an abnormal amount of a protein,
e.g. cancer.

XX Sequence 164 AA;
SQ

```
Query Match          27.8%; Score 532; DB 22; Length 164;  
Best Local Similarity 93.8%; Pred No. 7.7e+41;  
Matches 119; Conservative 1; Mismatches 22; Indels 0; Gaps 0;
```

CY	217	YRMEGGGFEGCLSDDECFHENTRETTSSSTTSRRFSLGYSIMADRPRTTQSGLATL	274
Dc	1	YRDMPGACSLKGSQMRYFTTPGHHPXSTRSPRSALGYSGDAXGRPITQLSELATAL	60
QY	227	AUASTPESSSHPTTFGTQHSSQTSMMSGVOSTPFTRDELFSJALOHALANGCPFLQS	336
Dc	61	AUASTPESSSHPTTFGTQHSSQTSMMSGVOSTPFINDLFSCALQALVAHQGPSLS	120
QY	337	QQQPQLQRDMXGIQDDLSLR	358
Dc	121	QQQPQLQRDMXGIQDDLSLR	142

RESULT 10
ID AGO13489 standard; Protein: 128 AA
AC AAO-3489;
DT 06-NOV-2001 (first entry)
DE Human polypeptide SEQ ID NO: 21461.
OS Homo sapiens.
PN WO200164835-A2.
PD 07-SEP-2001.
PX 26-FEB-2001; 2001WO-US04927.
PY XX
PR 28-FEB-2000; 2000US-0515126.
PP 18 MAY-2000; 2000US-0577409.
PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Drnanac RT;
XX WPI; 2001-514839/56.
XX N PSDB; AA193423.

Isolated nucleic acids and polypeptides, useful for preventing, diagnosing and treating e.g. leukaemia, inflammation and immune disorders -
Claim 20; SEQ ID NO 27381; 139pp + Sequence Listing; English.
The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00000-AAO13910); that exhibit activity relating to cytokine, cell proliferation or cell differentiation which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.
Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

Sequence 128 AA:
Query Match 26.5%; Score 566; DB 22; Length 128;
Best Local Similarity 96.4%; Pred No. 7.6e+39;
Matches 108; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

CY	1	MSSUDHLAVNKADQPLTKPSILRLPETELGEYSLGVSISFKQLIIAGKLQSVDPDEL	60
Dc	17	MSSUDHLAAAKLAQDLTKPSILRLPETELGEYSLGVSISFKQLIIAGKLQSVDPDEL	76
QY	61	IDLIYGCRKLKDCTLDVFYGIOPGSTVHVLRKSWEPECDKPSPVDKVAMRE	112
Dc	77	IDLIYGCRKLKDCTLCDFYGIOPGSTVHVLRKSWEPECDKPSPVDKEAMRD	128

RESULT 11
ID AAG02759 standard; Protein: 76 AA.
AC AAG02759;
DT 06-OCT-2000 (first entry)
DE Human secreted protein, SEQ ID NO: 684C.
KK Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
CS Homo sapiens.
XX EP1033403-A2.
PN EPI033403-A2.
PC C6-SEP-2000.
PX 21-FEB-2000; 2000EP-C2000510.
PR 26-FEB-1999; 99US-Q122487.
XX (GEST) GENSET.
PA Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
DR N-PSDB; AAC02765.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for diagnostic, forensic, gene therapy and chromosome mapping procedures -
Claim 13; SEQ ID 684C; 7ipp + CD-ROM; English.

CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or poly(A) RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA, because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 76 AA:
SQ

Query Match 18.2% Score 363; DB 21; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.4e-24;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSLSMDHMLAVKLAQDPLTPKSLRLPTELGEYSLGYSISFLKQLIAGKQESVPPPEL 60
Db : MSLSMDHMLAVKLAQDPLTPKSLRLPTELGEYSLGYSISFLKQLIAGKQESVPPPEL 60

QY 61 IDLIYCGPKLKDDDTL 76
Db 61 IDLIYCGPKLKDDDTL 76

RESULT 12
AAG30040
ID AAG30040 standard; Protein: 114 AA
AC AAG30040;
XX
XX 27-MAR-2001 (first entry)
XX
XX Human; breast cancer; ovarian cancer; cytostatic immunosuppressive;
XX neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
XX antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant;
XX antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
XX Addison's disease; allergy; autoimmune hemolytic anemia;
XX autoimmune thyroiditis; diabetes mellitus; chronic disease;
XX multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
XX cardiovascular disorder; wound healing; neurodegenerative disease;
XX
XX Homo sapiens.
XX
XX WO200005173-A1.
XX
XX 21-SEP-2000.
XX
XX 08-MAR-2000; 2000WO-US05881.
XX
XX 12 MAR-1999; 99US-0124270.
XX
XX (HUMAN) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI, 2000-61151/58.
XX
XX N-PSDB; AAF21683.
XX
XX New human breast and ovarian cancer associated gene sequences and the
XX polypeptides encoded by these genes, useful in the prevention,
XX treatment and diagnosis of cancer, immune disorders, cardiovascular
XX disorders and neurological diseases -
XX
XX Claim 11; Page 919; 1299pp; Eng. list.
XX
XX Sequences AAF1614 - AAF2003; present DNA sequences encoding human

CC proteins AAG58711 - AAB59128. The DNA and protein sequences are
CC associated with breast and ovarian cancer. Included in the invention are
CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
CC isolation and characterization of the DNA and protein sequences of the
CC invention. The breast and ovarian cancer associated DNA, protein, agonist
CC or antagonist sequences exhibit cytostatic, immunosuppressive;
CC neotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
CC antidiabetic; antiinflammatory; antitumor; antiparasitic; anticonvulsant;
CC antibacterial; antifungal; antiparasitic and cardiac activity. The
CC polynucleotide and protein sequences are used in the diagnosis of cancer,
CC particularly breast and ovarian cancer. The nucleic acid sequences,
CC proteins, agonists and antagonists may also be used in the diagnosis,
CC prevention and treatment of immune disorders e.g. Addison's disease,
CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
CC diabetes mellitus, chronic disease, multiple sclerosis, rheumatoid
CC arthritis and ulcerative colitis; cardiovascular disorders such as
CC myocardial ischaemia; wound healing; neurological diseases such as
CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 114 AA:
SQ

Query Match 17.3% Score 370; DB 21; Length 114;
Best Local Similarity 94.7%; Pred. No. 9.4e-23;
Matches 72; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSLSMDHMLAVKLAQDPLTPKSLRLPTELGEYSLGYSISFLKQLIAGKQESVPPPEL 60
Db : MSLSMDHMLAVKLAQDPLTPKSLRLPTELGEYSLGYSISFLKQLIAGKQESVPPPEL 60

QY 61 IDLIYCGPKLKDDDTL 76
Db 61 IDLIYCGPKLKDDDTL 97

RESULT 13
AAG30040
ID AAG30040 standard; Protein: 536 AA.
AC AAG30040;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 35843.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-030439.
XX
XX 25-FEB-1999; 99US-0121825.
XX
XX 05-MAR-1999; 99US-0123180.
XX
XX 09-MAR-1999; 99US-0123548.
XX
XX 23-MAR-1999; 99US-0125788.
XX
XX 25-MAR-1999; 99US-0126264.
XX
XX 29-MAR-1999; 99US-0126785.
XX
XX 01-APR-1999; 99US-0127462.
XX
XX 06-APR-1999; 99US-0128234.
XX
XX 08-APR-1999; 99US-0128714.
XX
XX 16-APR-1999; 99US-0129845.
XX
XX 19-APR-1999; 99US-0130077.
XX
XX 21-APR-1999; 99US-0130449.
XX
XX 23-APR-1999; 99US-0130510.
XX
XX 28-APR-1999; 99US-0131449.
XX
XX 30-APR-1999; 99US-0132048.
XX
XX 30-APR-1999; 99US-0132407.


```

PR 21-OCT-1999; 99US-0160768
PR 21-OCT-1999; 99US-0160770
PR 21-OCT-1999; 99US-0160814
PR 21-OCT-1999; 99US-0160815
PR 21-OCT-1999; 99US-0160980
PR 22-OCT-1999; 99US-0160981
PR 22-OCT-1999; 99US-0160989
PR 22-OCT-1999; 99US-0161404
PR 22-OCT-1999; 99US-0161405
PR 25-OCT-1999; 99US-0161406
PR 25-OCT-1999; 99US-0161359
PR 26-OCT-1999; 99US-0161360
PR 26-OCT-1999; 99US-0161361
PR 28-OCT-1999; 99US-0161320
PR 28-OCT-1999; 99US-0161392
PR 28-OCT-1999; 99US-0161393
PR 29-OCT-1999; 99US-0162142

Query Match 7.7% Score 164.5 EPR 21; length 536.
Best Local Similarity 24.0% EPR 21; No. 7.2e-05;
Matches 99; Conservative 49; Matches 140; Indels 125; Gaps 19;

QY 39 SISFLKQIAGLQSVDPFETGVYKGRKIKQCTIDFKGIQPGSTWIKRKWPPD 98
DE ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 TVSEFKELIA---QNSDFANQGVYKGRKIKQCTIDFKGIQPGSTWIKRKWPPD 95
QY 99 QNP-----EPVNVAAVREF-----RV 115
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
96 SAPAAAGNQTTPACVGSNDSSVLOGGSELPUSLPNPGDGNVAVGFPSSGPELQA 155
QY 116 LHAUHSSSSYE--AVERKANKSELDGIVATRKASQPIALGVLOCK DLFV 168
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
156 QCLAQKPMIREKNTPLAKNPPKPRAMVNNFQKRE-----DVRKHELHV 208
QY 169 FAPNMDLTLVPA HPAIVNAIIVHRSVAHAIWGTDSNMSKSSYKPMFQD 224
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
209 LNPRLRLCTLEARNPELMREV---NRNID RAN-SNLSMSEGNM 251
QY 225--LFEGLSDDDDEHENTRSTIGSTFSSFLASQNSHAGRPPTGCLLALALQA 279
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
252 LRMVENV--QPLAKNATTSEANNN--SSNFAALGNGNTTQOSTSNKISAPRAET 309
QY 280 STTESSEHTPTQCHSG--GIMWSTVQ--G-----ETPKN 315
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
310 GTFNAN--ELPWPWDAADOT--GPRNKKV--GPRNKKV--GPRNKKV 355
QY 316 NPLSAAAHACASLHPLG--GPRNKKV--GPRNKKV--GPRNKKV 379
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
367 QLSCILQNFVSKMGNVLSNLPVGGKNGNGLKMMVPRFGRMKNV 419

RESULT 14
AAG30039
ID AAG30039 standard; Protein; 577 AA.
AC
XX
AC AAG30039;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment 367 to 419; 1-843.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
CS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
FT 06-SEP-2000.
XX
IF 25-FEB-2000, 2500EP-0301439.
XX

```

```

PR 25-FEB-1999; 99US-0121825
PR 05-MAR-1999; 99US-0123180
PR 09-MAR-1999; 99US-0123548
PR 23-MAR-1999; 99US-0125788
PR 25-MAR-1999; 99US-0126264
PR 29-MAR-1999; 99US-0126785
PR 01-APR-1999; 99US-0127462
PR 06-APR-1999; 99US-0128234
PR 08-APR-1999; 99US-0128714
PR 16-APR-1999; 99US-0129845
PR 19-APR-1999; 99US-0130077
PR 21-APR-1999; 99US-0130449
PR 23-APR-1999; 99US-0130510
PR 23-APR-1999; 99US-0130891
PR 26-APR-1999; 99US-0131449
PR 30-APR-1999; 99US-0132048
PR 30-APR-1999; 99US-0132407
PR 04-MAY-1999; 99US-0132484
PR 05-MAY-1999; 99US-0132485
PR 06-MAY-1999; 99US-0132486
PR 06-MAY-1999; 99US-0132487
PR 07-MAY-1999; 99US-0132463
PR 11-MAY-1999; 99US-0134256
PR 14-MAY-1999; 99US-0134218
PR 14-MAY-1999; 99US-0134219
PR 14-MAY-1999; 99US-0134221
PR 14-MAY-1999; 99US-0134370
PR 18-MAY-1999; 99US-0134768
PR 19-MAY-1999; 99US-0134941
PR 20-MAY-1999; 99US-0135124
PR 21-MAY-1999; 99US-0135353
PR 24-MAY-1999; 99US-0135629
PR 25-MAY-1999; 99US-0136021
PR 27-MAY-1999; 99US-0136192
PR 28-MAY-1999; 99US-0136782
PR 01-JUN-1999; 99US-0137222
PR 03-JUN-1999; 99US-0137528
PR 04-JUN-1999; 99US-0137702
PR 07-JUN-1999; 99US-0137724
PR 08-JUN-1999; 99US-0138094
PR 10-JUN-1999; 99US-0138540
PR 10-JUN-1999; 99US-0138847
PR 14-JUN-1999; 99US-0139119
PR 16-JUN-1999; 99US-0139452
PR 16-JUN-1999; 99US-0139452
PR 17-JUN-1999; 99US-0139492
PR 18-JUN-1999; 99US-0139454
PR 18-JUN-1999; 99US-0139455
PR 18-JUN-1999; 99US-0139456
PR 18-JUN-1999; 99US-0139457
PR 18-JUN-1999; 99US-0139457
PR 18-JUN-1999; 99US-0139458
PR 18-JUN-1999; 99US-0139459
PR 18-JUN-1999; 99US-0139460
PR 18-JUN-1999; 99US-0139461
PR 18-JUN-1999; 99US-0139462
PR 18-JUN-1999; 99US-0139463
PR 18-JUN-1999; 99US-0139750
PR 18-JUN-1999; 99US-0139763
PR 21-JUN-1999; 99US-0139817
PR 22-JUN-1999; 99US-0139899
PR 23-JUN-1999; 99US-0140353
PR 23-JUN-1999; 99US-0140354
PR 24-JUN-1999; 99US-0140695
PR 28-JUN-1999; 99US-0140823
PR 29-JUN-1999; 99US-0140991
PR 30-JUN-1999; 99US-0141287
PR 01-JUL-1999; 99US-0141842
PR 01-JUL-1999; 99US-0142154
PR 02-JUL-1999; 99US-0142055
PR 06-JUL-1999; 99US-0142393
PR 08-JUL-1999; 99US-0142803
PR 09-JUL-1999; 99US-0142920
PR 12-JUL-1999; 99US-0142977

```

```
PR 13-JUL-1999; 99US-0143542;
PR 14-JUL-1999; 99US-0143624;
PR 15-JUL-1999; 99US-0144025;
PR 16-JUL-1999; 99US-0144285;
PR 19-JUL-1999; 99US-0144386;
PR 19-JUL-1999; 99US-0144325;
PR 19-JUL-1999; 99US-0144331;
PR 19-JUL-1999; 99US-0144332;
PR 19-JUL-1999; 99US-0144333;
PR 19-JUL-1999; 99US-0144334;
PR 19-JUL-1999; 99US-0144335;
PR 20-JUL-1999; 99US-0144352;
PR 20-JUL-1999; 99US-0144622;
PR 20-JUL-1999; 99US-0144684;
PR 21-JUL-1999; 99US-0144614;
PR 21-JUL-1999; 99US-0145065;
PR 21-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145088;
PR 22-JUL-1999; 99US-0145087;
PR 22-JUL-1999; 99US-0145089;
PR 22-JUL-1999; 99US-0145192;
PR 23-JUL-1999; 99US-0145145;
PR 23-JUL-1999; 99US-0145218;
PR 23-JUL-1999; 99US-0145224;
PR 26-JUL-1999; 99US-0145276;
PR 27-JUL-1999; 99US-0145913;
PR 27-JUL-1999; 99US-0145918;
PR 27-JUL-1999; 99US-0145919;
PR 28-JUL-1999; 99US-0145951;
PR 02-AUG-1999; 99US-0146166;
PR 02-AUG-1999; 99US-0146188;
PR 02-AUG-1999; 99US-0146389;
PR 03-AUG-1999; 99US-0147018;
PR 04-AUG-1999; 99US-0147224;
PR 05-AUG-1999; 99US-0147322;
PR 05-AUG-1999; 99US-0147192;
PR 06-AUG-1999; 99US-0147260;
PR 06-AUG-1999; 99US-0147303;
PR 06-AUG-1999; 99US-0147416;
PR 09-AUG-1999; 99US-0147493;
PR 09-AUG-1999; 99US-0147935;
PR 10-AUG-1999; 99US-0148171;
PR 11-AUG-1999; 99US-0148319;
PR 12-AUG-1999; 99US-0148341;
PR 13-AUG-1999; 99US-0148565;
PR 13-AUG-1999; 99US-0148684;
PR 16-AUG-1999; 99US-0149368;
PR 17-AUG-1999; 99US-0149115;
PR 18-AUG-1999; 99US-0149425;
PR 20-AUG-1999; 99US-0149222;
PR 20-AUG-1999; 99US-0149223;
PR 20-AUG-1999; 99US-0149929;
PR 23-AUG-1999; 99US-0149902;
PR 23-AUG-1999; 99US-0149930;
PR 25-AUG-1999; 99US-0150566;
PR 26-AUG-1999; 99US-0150884;
PR 27-AUG-1999; 99US-0151065;
PR 27-AUG-1999; 99US-0151066;
PR 27-AUG-1999; 99US-0151080;
PR 30-AUG-1999; 99US-0151303;
PR 31-AUG-1999; 99US-0151438;
PR 01-SEP-1999; 99US-0151935;
PR 01-SEP-1999; 99US-0152363;
PR 10-SEP-1999; 99US-0153070;
PR 13-SEP-1999; 99US-0153758;
PR 15-SEP-1999; 99US-0154018;
PR 16-SEP-1999; 99US-0154039;
PR 20-SEP-1999; 99US-0154779;
PR 22-SEP-1999; 99US-0155139;
PR 23-SEP-1999; 99US-0155486;
PR 24-SEP-1999; 99US-0155659;
PR 28-SEP-1999; 99US-0156458;
PR 29-SEP-1999; 99US-0156596;

PR 04-OCT-1999; 99US-0157117;
PR 05-OCT-1999; 99US-0157753;
PR 06-OCT-1999; 99US-0157865;
PR 07-OCT-1999; 99US-0158029;
PR 08-OCT-1999; 99US-0158232;
PR 12-OCT-1999; 99US-0158369;
PR 13-OCT-1999; 99US-0159293;
PR 13-OCT-1999; 99US-0159294;
PR 13-OCT-1999; 99US-0159295;
PR 14-OCT-1999; 99US-0159323;
PR 14-OCT-1999; 99US-0159330;
PR 14-OCT-1999; 99US-0159331;
PR 14-OCT-1999; 99US-0159637;
PR 14-OCT-1999; 99US-0159638;
PR 18-OCT-1999; 99US-0159584;
PR 21-OCT-1999; 99US-0162741;
PR 21-OCT-1999; 99US-0162767;
PR 21-OCT-1999; 99US-0160768;
PR 21-OCT-1999; 99US-0160770;
PR 21-OCT-1999; 99US-0160814;
PR 22-OCT-1999; 99US-0160815;
PR 22-OCT-1999; 99US-0160980;
PR 22-OCT-1999; 99US-0160981;
PR 22-OCT-1999; 99US-0160989;
PR 25-OCT-1999; 99US-0161404;
PR 25-OCT-1999; 99US-0161405;
PR 25-OCT-1999; 99US-0161406;
PR 26-OCT-1999; 99US-0161359;
PR 26-OCT-1999; 99US-0161360;
PR 26-OCT-1999; 99US-0161361;
PR 28-OCT-1999; 99US-0161920;
PR 28-OCT-1999; 99US-0161942;
PR 28-OCT-1999; 99US-0161993;
PR 29-OCT-1999; 99US-0162142;

Query Match 7.7%; Score 154.5; DB 21; Length 577;
Best Local Similarity 24.0%; Pred No Re-05;
Matches 99; Conservative 49; Mismatches 140; Indels 125; Gaps 19;

QY 33 S:SF:KQ:AGKLOESVPDPPEL:LIYCGRKLKDDQTLDFYGIQPGSTVHVLRKSWPEPD 98
DB 80 TVESFEKELIA:--QNSDVPANQOORLYKGRILKDDQTLISYGLQADHTVHMVGRGVFSSP 136
QY 99 QXP:-----EPDKVAAMREF:-----RV 115
DB 137 SAPAAVAGNQTTPACAVGNSNDSSNLJGGESLFPGLGFNPLAGGNAMAGLFGSLPLDEQA 196
QY 116 LHTALSSSSVRE:---AVFKMLSNKESLDCIIVATPGLSSDP:ALGVLOCK:--DLFSV 168
DB 191 QOCLAJNI:MIREFMMNTPAIQNLKXNPEFYKSMINNECMRE:-----JVERNPELGHV 249
QY 169 FADPNYL:LTLYPA:HPALVNAIVLVLSVAGSAPYPGTSDSSRSMPSYRDMPGGF:-- 224
DB 250 LNDPSIL:QTL:EAARNPELMREM:-----MRNTD:--RAM:--SNIESMPEGFNM 292
QY 225 ---LFEGLSDEDDPHNTRSTPSSSTPSSRP:--ASLGYSGAAGRPRTITQSELATALLA 279
DB 293 LRRMYENV:--QEPMLXATTTMSENAGNTSSNPFALLGNQGVTTQGSDTSNK:SAPNAET 350
QY 280 STPESSSH:PTGTGGRSSG:--TSPWSSGVOSG:-----TPITN 315
DB 351 GTPNAN: P:PNPWGATAGQTTAPGRTNAGLGGGLGGGLGGGLGMLGADSPGLGATPDAS 407
QY 316 DLFSQALOHALQASQPSLQSQWQ:-----POLQOOLRDMG:QDBELSIRP 359
DB 408 QL:SOILK:PAWSQWQSVLSNPQYMCMLMSLNPLQRLSMMLDMNPOLREMGNQP 459

RESULT 15
ABR69125
ID ABR69125 standard; Protein; 547 AA.
XX
AC ABR69125;
```


GenCore version 5.1.6
Copyright (c) 1993 2003 CompuGen Ltd.

OM protein - protein search, using sw modli

Run on: November 12, 2003, 06:18:19 / Search time 30 Seconds
without alignments
234 / 240 Million cell: updates/sec

Title: US-09-735-251-4
Perfect score: 2133
Sequence: 1 MS:SDMR:AVKLADQLTPR... / SGPYGYCWEADYKVERSEV 410

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 644079 seqs, 171749292 residues
Total number of hits satisfying chosen parameters: 644079

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications: AA*
1: /cgn2_6/ptodata/1/pubaa/US01_PUBCMB.pep.
2: /cgn2_6/ptodata/1/pubaa/US02_PUBCMB.pep.
3: /cgn2_6/ptodata/1/pubaa/US06_NEW_PUBCMB.pep.
4: /cgn2_6/ptodata/1/pubaa/US06_PUBCMB.pep.
5: /cgn2_6/ptodata/1/pubaa/US07_NEW_PUBCMB.pep.
6: /cgn2_6/ptodata/1/pubaa/US07_PUBCMB.pep.
7: /cgn2_6/ptodata/1/pubaa/US08_NEW_PUBCMB.pep.
8: /cgn2_6/ptodata/1/pubaa/US08_PUBCMB.pep.
9: /cgn2_6/ptodata/1/pubaa/US09_PUBCMB.pep.
10: /cgn2_6/ptodata/1/pubaa/US09_PUBCMB.pep.
11: /cgn2_6/ptodata/1/pubaa/US09_PUBCMB.pep.
12: /cgn2_6/ptodata/1/pubaa/US09_NEW_PUBCMB.pep.
13: /cgn2_6/ptodata/1/pubaa/US10_PUBCMB.pep.
14: /cgn2_6/ptodata/1/pubaa/US10_PUBCMB.pep.
15: /cgn2_6/ptodata/1/pubaa/US10_PUBCMB.pep.
16: /cgn2_6/ptodata/1/pubaa/US10_PUBCMB.pep.
17: /cgn2_6/ptodata/1/pubaa/US10_NEW_PUBCMB.pep.
18: /cgn2_6/ptodata/1/pubaa/US10_PUBCMB.pep.

ALIGNMENTS				RESULT 1			
Result No.	Score	Query Match	Length DB ID	Description	Sequence	Length	DB ID
1	1834	86.0	380	US-09-890-688-86	Sequence 86, Appl	1750	15
2	1834	86.0	420	US-09-729-674-172	Sequence 172, Appl	1203	15
3	1834	17.3	114	US-10-102-906-438	Sequence 488, Appl	1259	15
4	152.5	7.1	624	US-10-146-474-47	Sequence 47, Appl	1210	15
5	141.5	6.6	589	US-10-293-000-4	Sequence 4, Appl	778	12
6	141.5	6.6	595	US-10-293-000-2	Sequence 2, Appl	778	12
7	141	6.6	528	US-09-840-746-2	Sequence 20, Appl	524	15
8	135.5	6.4	1162	US-09-745-083-34	Sequence 34, Appl	1138	15
9	134	6.3	655	US-10-601-612A-2	Sequence 2, Appl	1063	12
10	134	6.3	655	US-10-054-681-29	Sequence 29, Appl	780	12
11	128	6.0	1134	US-09-816-392-16	Sequence 16, Appl	1367	10
12	127.5	6.0	881	US-09-816-860A-2	Sequence 2, Appl	1056	14
13	126	5.9	2971	US-10-146-473-52	Sequence 50, Appl	1237	10
14	124.5	5.8	841	US-10-029-306-320-4	Sequence 32014, A	2011	15
15	124.5	5.8	1714	US-10-128-714-3176	Sequence 3176, Ap	345	9

Pred. No. is the number of predicted proteins in the database that have a score greater than or equal to the score of the hit. The number of hits is listed and is derived by analysis of the total database distribution.

SUMMARY:

US-09-890-688-86
; Sequence 86, Application: US/09890688
; Publication No. US20030144475A1
; GENERAL INFORMATION:
; APPLICANT: Setahi KATO
; APPLICANT: Chikashi EGUCHI
; APPLICANT: Minoru SAKI
; TITLE OF INVENTION: Human Proteins and cDNAs thereof
; FILE REFERENCE: 2001-1102A/WMC/C0653
; CURRENT APPLICATION NUMBER: US/09/890,688
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: JP 11-346863
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 11-34684
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: JP 2000-31062
; PRIOR FILING DATE: 2000-02-08
; PRIOR APPLICATION NUMBER: JP 2000-34091
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-34090
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: JP 2000-35829
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-35899
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: JP 2000-71161
; PRIOR FILING DATE: 2000-03-14
; PRIOR APPLICATION NUMBER: JP 2000-160851
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 86
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-890-688-86

Query Match 86.0%, Score 1834, DB 12, Length 380,
 Best Local Similarity 100.0%, Pred. No. 1, 1e-134,
 Matches 358, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 60
 DB 1 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 60

QY 61 IDLIYCGRKLKDDQTLDFYGCSTVHVLRKKSWPEPQKLEPVKVAAMBEFRVLHTAL 120
 DB 61 IDLIYCGRKLKDDQTLDFYGCSTVHVLRKKSWPEPQKLEPVKVAAMBEFRVLHTAL 120

QY 121 HSSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPPIALGVLDQKDLFSVFADPNMLDTLVP 180
 DB 121 HSSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPPIALGVLDQKDLFSVFADPNMLDTLVP 180

QY 181 APPALVNAIVLVLSVAGSAPVPGTSSSRSPSSSYRDMPCGFLFEGLSDEDDPHNT 240
 DB 181 APPALVNAIVLVLSVAGSAPVPGTSSSRSPSSSYRDMPCGFLFEGLSDEDDPHNT 240

QY 241 RSTPSSSTPSRRPASLGYSGAAGPRPTTQSELATATLALASTPSSSHTPTPTQGHSSGT 300
 DB 241 RSTPSSSTPSRRPASLGYSGAAGPRPTTQSELATATLALASTPSSSHTPTPTQGHSSGT 300

QY 301 SPSSSGVCGSTPTINDLFSQALQHALQASQFSLOSOWOPQQLQURDMGICQDELSLR 358
 DB 301 SPSSSGVCGSTPTINDLFSQALQHALQASQFSLOSOWOPQQLQURDMGICQDELSLR 358

RESULT 2

US-09-729-674-172
 ; Sequence 172, Application US/09129674
 ; Publication No. US20010339335A;
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.
 ; APPLICANT: LaVallie, Edward R.
 ; APPLICANT: Collins-Racie, Lisa A.
 ; APPLICANT: Evans, Cheryl
 ; APPLICANT: Moberg, David
 ; APPLICANT: Treacy, Maurice
 ; APPLICANT: Agostino, Michael J.
 ; APPLICANT: Steinginger II, Robert J.
 ; APPLICANT: Spaulding, Vikki
 ; APPLICANT: Wong, Gordon G.
 ; APPLICANT: Clark, Hilary
 ; APPLICANT: Fichtel, Kim
 ; APPLICANT: Genetics International
 ; TITLE OF INVENTION: SECRETED RE COMBINATION
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09-729-674
 ; PRIOR FILING DATE: 2000-03-10
 ; PRIOR APPLICATION NUMBER: 09/573,410
 ; NUMBER OF SEQ ID NOS: 283
 ; SOFTWARE: Patentin Ver. 2.3
 ; SEQ ID NO 172
 ; LENGTH: 420
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-729-674-172

Query Match 86.0%, Score 1834, DB 9, Length 410,
 Best Local Similarity 100.0%, Pred. No. 1, 1e-134,
 Matches 358, Conservative 0, Mismatches 0, Indels 0, Gaps 0;

QY 1 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 60
 DB 41 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 100

QY 61 IDLIYCGRKLKDDQTLDFYGCSTVHVLRKKSWPEPQKLEPVKVAAMBEFRVLHTAL 120
 DB 101 IDLIYCGRKLKDDQTLDFYGCSTVHVLRKKSWPEPQKLEPVKVAAMBEFRVLHTAL 160

QY 121 HSSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPPIALGVLDQKDLFSVFADPNMLDTLVP 180
 DB 161 HSSSSYREAVFKMLSNKESLDQIIIVATPGLSSDPPIALGVLDQKDLFSVFADPNMLDTLVP 220

QY 181 APPALVNAIVLVLSVAGSAPVPGTSSSRSPSSSYRDMPCGFLFEGLSDEDDPHNT 240
 DB 221 APPALVNAIVLVLSVAGSAPVPGTSSSRSPSSSYRDMPCGFLFEGLSDEDDPHNT 280

QY 241 RSTPSSSTPSRRPASLGYSGAAGPRPTTQSELATATLALASTPSSSHTPTPTQGHSSGT 300
 DB 281 RSTPSSSTPSRRPASLGYSGAAGPRPTTQSELATATLALASTPSSSHTPTPTQGHSSGT 340

QY 301 SPSSSGVCGSTPTINDLFSQALQHALQASQFSLOSOWOPQQLQURDMGICQDELSLR 358
 DB 341 SPSSSGVCGSTPTINDLFSQALQHALQASQFSLOSOWOPQQLQURDMGICQDELSLR 398

RESULT 3

US-10-102-806-488
 ; Sequence 488, Application US/10102806
 ; Publication No. US20030034421A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA/031PCL
 ; CURRENT APPLICATION NUMBER: US/10-102-806
 ; CURRENT FILING DATE: 2002-03-22
 ; PRIOR APPLICATION NUMBER: 09/925,298
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05881
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 63/124,270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 846
 ; SOFTWARE: Patentin Ver. 2.3
 ; SEQ ID NO 488
 ; LENGTH: 114
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (95)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (111)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; NAME/KEY: SITE
 ; LOCATION: (111)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 ; US-10-102-806-488

Query Match 17.3%, Score 370, DB 15, Length 114,
 Best Local Similarity 94.7%, Pred. No. 2, 3e-21,
 Matches 72, Conservative 2, Mismatches 2, Indels 0, Gaps 0;

QY 1 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 60
 DB 22 MSLSDWHLAVKLADQPLTPKSIILPLPETELGEVSLGGYSISFLKQIAGKLOESVDPDEL 81

QY 61 IDLIYCGRKLKDDQTL 76
 DB 82 IDLIYCGRKLKDDQTL 97

RESULT 4

US-10-146-473-47
 ; Sequence 47, Application US/10146473
 ; Publication No. US2003010888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Scanlan, Matthew
 ; APPLICANT: Gout, Ivan
 ; APPLICANT: Stockert, Elisabeth

APPLICANT: Gure, Ali
 APPLICANT: Chen, Yao-Teung
 APPLICANT: Old, Lloyd
 TITLE OF INVENTION: Breast Cancer Antigens
 FILE REFERENCE: 600461/70130 (SRV)
 CURRENT APPLICATION NUMBER: US/10/746,473
 CURRENT FILING DATE: 2002-05-17
 PRIOR APPLICATION NUMBER: US 60/299,150
 PRIOR FILING DATE: 2001-05-15
 NUMBER OF SEQ ID NOS: 82
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 47
 LENGTH: 624
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-293-000-4

Query Match 7.1%; Score 152.5; DB 15; Length 624;
 Best Local Similarity 19.8%; Pred. No. 0.0221;
 Matches 97; Conservative 70; Mismatches 17; Indels 147; Gaps 19;

QY 13 ADQPLTKS: LRLPETELGEYSL-GGYS:SELKQLIAGKLOESVDPDLIIYCG 67
 DB 24 AAAPAEKIIKVTYKTPK-EKEEFAPVENSVOQFKKESIKRFXSHT---DOLVLIAG 82
 QY 69 KKKKQDQTLDFYGIQPGSTVHLRKSWEPPDKPEPVKVAAMREFRVLRLTAHSSSSYR 127
 DB 80 ILKKQDQTLDFYGIQPGSTVHLRKSWEPPDKPEPVKVAAMREFRVLRLTAHSSSSYR 127
 QY 128 EAVFKMLSKESLQDIIIVATP-----GLSS-----DPI 155
 DB 131 SNSTPSTNSPFGSLGSLGAGLSSGLSLGNTTNFSELSQSQMQLSNPEM 185
 QY 156 ALGVLOKCLFVFPADPNMLDILVPAHPALVNA:VL---VLHVSAGSAMPCTDSSSRSM 212
 DB 191 VQMLSNPDLMLQIMANPQV-QDILKRNPEISHLNPNQIMOTIMOTIELARN- 244
 QY 209 SRSPSSSYRMPGGF-LFEGSLGSLGAGLSSGLSLGNTTNFSELSQSQMQLSNPEM 185
 DB 250 NQDIALSNLESIPGGYNALRRPYDIDFPM:MAVJEQFGSH---ASVSSSSSGIC 306
 QY 264 -----PPITQSELATLALASTPSSSHT-----P 289
 DB 307 PRTYERKPLNPWAPPATSSS ATSTSTTSJGNSGNSGNTGNTYMAVVAIES 365
 QY 290 TPGIQQ ---MAVVAIESGNTGNTYMAVVAIES 365
 DB 366 TPGIQQ---MAVVAIESGNTGNTYMAVVAIES 365
 QY 338 WCPHGLQRLNGLDDELSLRPRCPVPVGTSGKOPWSSSLLEPHELP 383
 DB 426 MRPGD-----PFLQOMQQLTFLQOMQ 461
 QY 338 WEALMKVPPSL 408
 DB 452 MQALMQIQOGL 462

RESULT 5
 US-10-293-000-4
 Sequence 4, Application US/10293000
 Publication No. US20030175278A1
 GENERAL INFORMATION:
 APPLICANT: Monteiro, Mervyn J.
 APPLICANT: Mah, Alex L.
 APPLICANT: Perry, George
 APPLICANT: Smith, Mark A.
 TITLE OF INVENTION: UBIQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
 FILE REFERENCE: 4115-175
 CURRENT APPLICATION NUMBER: US/10/293,000
 CURRENT FILING DATE: 2003-04-01
 PRIOR APPLICATION NUMBER: US 60/238,549
 PRIOR FILING DATE: 2001-11-13

NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 4
 LENGTH: 589
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-293-000-4

Query Match 6.6%; Score 141.5; DB 12; Length 589;
 Best Local Similarity 18.6%; Pred. No. 0.014;
 Matches 94; Conservative 81; Mismatches 169; Indels 161; Gaps 19;

QY 9 AVKADCPDTPKSLRLPETELGEYSL-GGYS:SELKQLIAGKLOESVDPDLIIYCG 67
 DB 27 AAASAEPRXIKVTYKTPK-EKEEFAPVENSVOQFKKESIKRFXSHT---DOLVLIAG 82
 QY 68 RKKKDQDTEFYGIQPGSTVHLRKSWEPPDKPEPVKVAAMREFRVLRLTAHSSSSYR 127
 DB 83 KILKQDQTLDFYGIQPGSTVHLRKSWEPPDKPEPVKVAAMREFRVLRLTAHSSSSYR 127
 QY 128 EAVFKMLSKESLQDIIIVATP-----GLSS-----DPI 155
 DB 126 NVTTSSTPNSSTSSGATSNPFGSLGSLGAGLSSGLSLGNTTNFSELSQSQMQLSNPEM 185
 QY 156 ALGVLOKCLFVFPADPNMLDILVPAHPALVNA:VL---VLHVSAGSAMPCTDSSSRSM 212
 DB 196 MVOIXENPFGVMSLNPDLNQLIMANPQV-QDILKRNPEISHLNPNQIMOTIMOTIELARN- 244
 QY 213 PS-----SSYRMPGGF-----LFEG-----LSDDEDFHNTSTPSSSTP 249
 DB 245 PAMQEMMNQDRALSNLESIPGGYNALRRPYDIDFPM:MAVJEQFGSH---ASVSSSSSGIC 306
 QY 250 SRSPASLGYSGAGPPP-----ITQSELATLALASTPSSSHTPTPTQGHSSGTS 303
 DB 305 SGEQSFSTRNRPDLPNPWAPPATSSS ATSTSTTSJGNSGNSGNTGNTYMAVVAIES 365
 QY 304 SSVQVS-----GTP:TNDFLS-----CALQHALQA----- 328
 DB 364 VPGVGSMTXNTGMSLQDIIIVATP-----P 289
 QY 329 -SQPSLSQSQWQPQ-----LQQLRDMGIQDDELSLRPRCPVPVGTSGKOPWSSSLLEPHELP 383
 DB 424 FAGNPOLCEQMQLTFLQOMQ-----PDTLSMNSPRAHQ 461
 QY 364 A-----SPEPPASCRGYCP 397
 DB 452 ALQIQOGLTATEAPGLTPFTP 496

RESULT 6
 US-10-293-000-2
 Sequence 2, Application US/10293000
 Publication No. US20030175278A1
 GENERAL INFORMATION:
 APPLICANT: Monteiro, Mervyn J.
 APPLICANT: Mah, Alex L.
 APPLICANT: Perry, George
 APPLICANT: Smith, Mark A.
 TITLE OF INVENTION: UBIQUILIN, A PRESENILIN INTERACTOR AND METHODS OF USING SAME
 FILE REFERENCE: 4115-175
 CURRENT APPLICATION NUMBER: US/10/293,000
 CURRENT FILING DATE: 2003-04-01
 PRIOR APPLICATION NUMBER: US 60/238,549
 PRIOR FILING DATE: 2001-11-13
 NUMBER OF SEQ ID NOS: 8
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 2
 LENGTH: 595
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-293-000-2

Query Match 6.6%; Score 141.5; DB 12; Length 595;


```

237 PIPSTTSVQPSSSSAPTTSATSVQPSSSSPTPIPTTSVQPSSSSAPTTSATSVQPS 290
224 FLFELGSDDEDHFHENTRSTPSSSTPSSRPASLG---YSGAAGRPPIITQSELATA.LALAS 280
297 SSSPSTTSITSVQPSSSSSSPTTSITSVQPSSSSAPTTSATSVQPSSSSPTPISTISV 356
281 TPESSSH:PTGCT---QCHSSGCTPMESGV-----OSGTPI-T 314
357 QSSSSSSSPTTSVQPSSSGSAPTTSATSVQPSSSSVPTTSATSVRSSSSSPTPIPT 416
315 NDLFSQALQHALQASQPSLGSQWQPLQQLRDMGIODELSLRPCR-----PPVGT:SKO 369
412 TS-SVQPSSSSVPTTSATSVQTSSSSSTPIPTTSVQPSSSSAPTTSATSVQPSSSSP 476
370 PMSSSSL:EPHELPRASP-----EPPASCRGICPFWAL:KRVPS 407
477 PISSTTSVQPSSSSPTTSITSVQPSSS--GSAPTTSATSVQPS 519

RESULT 8
US-09-745-008-34
: Sequence 34, Application US/09745008
: Patent No. US2006037667A1
: GENERAL INFORMATION:
: APPLICANT: Chuenkova, Marina
: APPLICANT: Pereira, Mircio A.
: TITLE OF INVENTION: T. Cruzi-Derived Neurotrophic Agents and
: TITLE OF INVENTION: Methods of Use Therefor
: FILE REFERENCE: 122/1028-001
: CURRENT APPLICATION NUMBER: US/09/745,008
: PRIOR FILING DATE: 2006-12-20
: PRIOR APPLICATION NUMBER: US 60/172,881
: PRIOR FILING DATE: 1999-12-20
: NUMBER OF SEQ ID NOS: 34
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 1162
: TYPE: ppt
: ORGANISM: Trypanosoma cruzi
US-09-745-008-34

Query Match 6.4%; Score 135.5; DB 10; Length 1162;
Best Local similarity 21.8%; Pred No. 0.1;
Matches 98; Conservative 51; Mismatches 189; Indels 111; Gaps 18

QY 6 WHLAVKADOP:TPRSILRLPTELGEYSIGGYS:FLKOLLACKLOESVDPDEL:D--- 62
DQ 499 YHGVTHMAN -----KIGSVYIDG-----ELLESGGTVVFDGRTPDISH 537
QY 43 -----LIYGRKKKXDDOTLDFYGIQP--GSTVHLRKSMPED 98
DQ 518 FYVGQYKGFEMPTISHVTWNVLLYNPQLNTEETRLFLSQDLIGTEAHMDSSSDSSAH 597
QY 49 OKPE EVCKVAAKRFVRLVHLTHALSSSYREAVFKMLSNKESLDIIVATPG:SSPIAL 157
DQ 598 STSTFALUSAHST:STTPVDSAHSTSTPADSSAHGCTPTPVDSSAHG:FSTPADSSAH 657
QY 158 GYLQKQKGF:SVFADENM LDTLVPAHVALYNAIVL:VHVAGAPMPGCTDSSSRVPS-- 214
DQ 652 GTPSTHY:SAHSTSTPVDSSAHSTFS-----TPVDSSAHGAPSTPADSSAHGTPSTP 711
QY 215 --SSVEKPGGFLFGLSCDDEDDHPN:TRSTPSSSTPSSRPASL:GYSGAAGRPPIITQSEL 272
DQ 712 VDSSAHG:P-----STPADSSAHSTPTPADSSAHSTPSTPADS:SAHSTFS:PVD- 761
QY 273 ATALALASTPSSSHTP:PGT---QCHSSGCTSPMSSGVOS--GTPTINDLFSCALQHAL 326
DQ 762 SNAHCTPTPADSSAHSTPS:PADSSAHG:PTSTPVDSSAHSTSTPTVDS----- 810
QY 327 QASGQF:SLGSQWQPLQQLRDMGIODELSLRPCR-----PVTGSKOPWSSSLLEPHEL 382
DQ 811 SAHTF:STP-----VDSAHSTPTPVDSSAHG:TPSTPVDSS-----AHST 851

```

DB	712	VDSAHG-P-----STPADSSAHSTPADSSAHSTPADS-SAHSTPS-PVD- 761
QY	273	ATALALATPESSSHTP-PG- ---QCRSGTSPMSSGVQS---GTITNDLFSQALQHAL 326
DB	762	SSAHGTPSTPADSSAHSTPS-PADSSAHGTPSTPVDSSAHSTPTSPVDS- 810
QY	327	QASGQFSLQSQMCPQLQCLRMGMIDGELSURPCP- ---PVGTQKQPMSSSLLEPHEL 382
DB	911	SAHSTPS-T- ---VCGSSAHSTPTPVDSSAHGTPSTTPVDS- ---AHST 951

QY 383 PASP-EPPASCRGVCWEALMKVPPSLPV 410
Db 852 PSTPAUSSAHTSTPADSSAHGTPTSPV 899

RESULT 9
US-10-001-632A-2
; Sequence 2, Application US/1000-612A
; Publication No. US20020151492A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Carrell C.
; APPLICANT: Feinhaus, Andrew L.
; APPLICANT: Holderman, Susan D.
; TITLE OF INVENTION: Tesis Specific Protein
; FILE REFERENCE: 99-1701
; CURRENT APPLICATION NUMBER: US/1000-612A
; PRIOR FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: 60/128,210
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/156,940
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 69/541,9190
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Vers: 3.0
; SEQ ID NO 2
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-001-632A-2

Query Match 6.38; Score 134; DB 15; Length 655;
Best Local Similarity 19.48; Pred. No. 0.062;
Matches 86; Conservative 61; Mismatches 111; Indels 146; Gaps 16;

QY 39 SIFLKQIAGKLGESVDPPELIDLYCORKLKDDQTLDFYQ:QPGSTVHVLRK----- 92
Db 42 TIOQKEIS---QRFKAHPDQVLIFAGKILKDPDSLAQCGVRDGLTVHVLKIQHRAH 98
QY 93 -----SWPEPQCKPEP-----VDKVAAMREFRVLHTAL--HSSSSY 126
Db 99 GNECPAASVPTQGPSFQSPGLQPSS:YPADGPPAPFSLGLTGLSLGLAYRFPDPQSSLM 158
QY 127 REAVFKM:SNKESLDQIIIVATPGLSDDIALGVLCQKDLFS--VFADPNMUTL----- 178
Db 159 RQHV---SVPEFVQLI---DDPFGLLSNTGLVRQLVLDNPHMQCLIQHNPEI 207
QY 179 --VPAHPALNNAIVLVHVSAGSAPKPGTSSSRMPSSSYDMPPGF-----LFEGLSD 231
Db 208 GHILNPEIMVROTLEFLNPNAMQEV---RSQDRVLSNLESIPGGYNVJCTWYTDMD 263
QY 232 D-----ECCFHN---TRSTSSSTPSRPSASLGYSAGAPRPITQSELATALALATPE 283
Db 264 PMLNAVQOFGGNPFATATTNATTTSCPSRX-----E 297
QY 284 SSSHTPTGTGHSSTGSPMSSGVQSGTPITNDLFS-----QA 321
Db 299 NCDPLNPNWSTH--GGSGRQGGQGGQDAF:SNRFPNFLGI:RLVYDYLQQLCHENPQS 355
QY 322 LQHALQASGQPSLQSQWQPCQLRLMGLQDELRLRCRPPVGTSGKPFWSSSLLEPHE 381
Db 356 LGTYLQGTASALSQSQ-----EPPSVNRPVPPSPSSQEPQS 392
QY 382 LPASPEPPASCRGVCWEALMKVP 405
Db 393 GQPLPEESVAIKGRSSCPAFRLYP 416

RESULT 10
US-10-054-683-29
; Sequence 29, Application US/10054683
; Publication No. US2003004481A1
; GENERAL INFORMATION:

APPLICANT: Old, Lloyd J.
APPLICANT: Scanlan, Matthew J.
APPLICANT: Chen, Yao-Tseng
TITLE OF INVENTION: Cancer-Testis Antigens
FILE REFERENCE: LC461/7125 (JRV)
CURRENT APPLICATION NUMBER: US/10/054,683
CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/280,718
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/285,154
PRIOR FILING DATE: 2001-04-20
PRIOR APPLICATION NUMBER: US 60/327,432
PRIOR FILING DATE: 2001-10-05
NUMBER OF SEQ ID NOS: 37
SOFTWARE: PatentIn version 3.1
SEQ ID NO 29
LENGTH: 655
TYPE: PRT
ORGANISM: Homo sapiens
US-10-054-683-29

Query Match 6.38; Score 134; DB 15; Length 655;
Best Local Similarity 19.48; Pred. No. 0.062;
Matches 86; Conservative 61; Mismatches 151; Indels 146; Gaps 16;

QY 39 SIFLKQIAGKLGESVDPPELIDLYCORKLKDDQTLDFYQ:QPGSTVHVLRK----- 92
Db 42 TIOQKEIS---QRFKAHPDQVLIFAGKILKDPDSLAQCGVRDGLTVHVLKIQHRAH 98
QY 93 -----SWPEPQCKPEP-----VDKVAAMREFRVLHTAL--HSSSSY 126
Db 99 GNECPAASVPTQGPSFQSPGLQPSS:YPADGPPAPFSLGLTGLSLGLAYRFPDPQSSLM 158
QY 127 REAVFKM:SNKESLDQIIIVATPGLSDDIALGVLCQKDLFS--VFADPNMUTL----- 178
Db 159 RQHV---SVPEFVQLI---DDPFGLLSNTGLVRQLVLDNPHMQCLIQHNPEI 207
QY 179 --VPAHPALNNAIVLVHVSAGSAPKPGTSSSRMPSSSYDMPPGF-----LFEGLSD 231
Db 208 GHILNPEIMVROTLEFLNPNAMQEV---RSQDRVLSNLESIPGGYNVJCTWYTDMD 263
QY 232 D-----ECCFHN---TRSTSSSTPSRPSASLGYSAGAPRPITQSELATALALATPE 283
Db 264 PMLNAVQOFGGNPFATATTNATTTSCPSRX-----E 297
QY 284 SSSHTPTGTGHSSTGSPMSSGVQSGTPITNDLFS-----QA 321
Db 299 NCDPLNPNWSTH--GGSGRQGGQGGQDAF:SNRFPNFLGI:RLVYDYLQQLCHENPQS 355
QY 322 LQHALQASGQPSLQSQWQPCQLRLMGLQDELRLRCRPPVGTSGKPFWSSSLLEPHE 381
Db 356 LGTYLQGTASALSQSQ-----EPPSVNRPVPPSPSSQEPQS 392
QY 382 LPASPEPPASCRGVCWEALMKVP 405
Db 393 GQPLPEESVAIKGRSSCPAFRLYP 416

RESULT 11
US-09-936-392-16
; Sequence 16, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides,
; FILE REFERENCE: PTO2001
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1995-10-15

PRIOR APPLICATION NUMBER: 60/165,914
 PRIOR FILING DATE: 1999-11-17
 PRIOR APPLICATION NUMBER: 60/189,027
 PRIOR FILING DATE: 2002-03-14
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 16
 LENGTH: 1134
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-816-860A-2

Query Match 6.0%; Score 128, DB 9; Length 881
 Best Local Similarity 24.7%; Pred. No. 2.1;
 Matches 69; Conservative 31; Mismatches 38; Indels 96; Gaps 14;

QY 195 SVAGSAPVCTGS---SSSVSSSVYRMPGPHLEGLSDDEDFHNTPTSTPSSTPS 251
 DB 681 SLSSSSLSGSPHGLPARS-PHSVSTSTPS-AVLGASSQS ---SSPASTTNS 730
 QY 252 RPASLYSGAAGRPPTOSLATALA-----ALASTESSSHPTPTPT 293
 DB 731 -PAS---SASHH-RPSTLHSLSP-HROVESAPKAGNIFNP-LAHITSTQAPPP-L 785
 QY 294 CGHSSST-----SPMSSTVQS-----HPTNHLQSQ 320
 DB 786 PHTVSSHTTQFPKILNSPPVVEFPKSAEPREPRLKSVQSAEKALGASAKAKGA 845
 QY 321 ALQHALLCAGQPSLQSQWQ-----POCCALMMHQTCELSL 357
 DB 846 LRKHSLEV-GHPDFKDFCELALHSLARSUSHTPEVEGLAPROQVAVRQARQSPSLD 904
 QY 358 RPCRPP-PUTSQPSSSLDEHPELHSPSPFASCPG 344
 DB 905 -GADPLLEGASRPVSSKESKESQGAETHPKATTPG 944

RESULT 13
 US-09-816-860A-2
 Sequence 2, Application US/99814660A
 Patent No. US200001651A1
 GENERAL INFORMATION:
 APPLICANT: Meyers, Rachel
 TITLE OF INVENTION: 26649, A NO. US20020208, Eliezer Human Cytosine Activating Molecule a
 TITLE OF INVENTION: Uses for the
 FILE REFERENCE: MN: 133
 CURRENT APPLICATION NUMBER: US 09-816-860A
 PRIOR APPLICATION NUMBER: 60/165,914
 PRIOR FILING DATE: 1999-11-17
 PRIOR FILING DATE: 2002-03-14
 NUMBER OF SEQ ID NOS: 14
 SOFTWARE: Patent In Ver. 2.0
 SEQ ID NO 2
 LENGTH: 881
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-816-860A-2

Query Match 6.0%; Score 127, DB 9; Length 881
 Best Local Similarity 25.3%; Pred. No. 2.1;
 Matches 77; Conservative 27; Mismatches 50; Indels 94; Gaps 14;

QY 158 GVLDKDLFSVF---ADPNKDLTVAHPALQALNAIVLHVSAGHAPVWGTSSSRMPS 214
 DB 512 GT-LNRKHISAFQPPPLPTQSGT-WP-----ATPEFH CSSRASS 552
 QY 215 SSVYRMP--GGFLPEG-SCSDDFHNTPTSTPSSTPS-----SP 252
 DB 553 SGGTVPSSAGILLEGQSPG-DSPKPKDQVSAAVPAKPAPKNSQASLQNFCAAASH 611
 QY 253 PASLGY-SGAGRPPTCCSELATALALASTESSSHPTPTPTGOG-HSSGTWPSGQS 309
 DB 612 QLSMSQPHNAAGSPHT---DRFAVKKPAPAPKPKGNPPPGIPGGQSSSGTCHPPPSLP 669

QY 310 GTPITNDFESQALQHALLCAGQPSLQSQWQPMQLOLRDMGIQDDLSLRPCRPPVGT-SKQ 369
 DB 669 KPTTRSP-SPPTQHTGPPGQPSAPSLSAPRRY-----SSLSPICAPNHPPPQ 717
 QY 370 PWSSSL-----LEEP-HELSPASPEPPASCRGCPWEALMKVP 405
 DB 718 PPTCAITLMTKPNQSGPPNPMALPSEHGLEBQPSHTPTPTTPTST-----PPLGKQN 770
 QY 406 PSLP 409
 DB 771 PSLP 714

RESULT 13
 US-10-146-473-50
 Sequence 50, Application US/10146473
 Publication No. US20030108888A1
 GENERAL INFORMATION:
 APPLICANT: Scania, Matthew
 APPLICANT: Golt, Ivan
 APPLICANT: Stockert, Elisabeth
 APPLICANT: Gure, Ali
 APPLICANT: Chen, Yao Tseng
 APPLICANT: Old, Lloyd
 TITLE OF INVENTION: Breast Cancer Antigens
 FILE REFERENCE: US46170130(JPV)
 CURRENT APPLICATION NUMBER: US/10/146,473
 CURRENT FILING DATE: 2002-05-15
 PRIOR APPLICATION NUMBER: US 60/291,150
 PRIOR FILING DATE: 2001-05-15
 NUMBER OF SEQ ID NOS: 82
 SOFTWARE: Patent In version 3.0
 SEQ ID NO 50
 LENGTH: 2971
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-146-473-50

Query Match 5.9%; Score 126; DB 15; Length 2971;
 Best Local Similarity 22.7%; Pred. No. 2.1;
 Matches 114; Conservative 49; Mismatches 178; Indels 162; Gaps 23;

QY 12 LABQPLTPKSL-----RLPETELGEYSLGGYISFLKQLIAGKLOESVDP 58
 DB 1268 LAPAPRPSGSLPAVLNPRPTLTGRLPTTLOTARAPMTTTLVRPL--KLVS-PSP 1124
 QY 59 ELIQLYCGKRLKDDOTLQFYGIOP---GSTVHVLRKSWPEPDQKPEPVKVAAMREFRV 115
 DB 1125 EV-----SASAPGAAPLTSSPLHV-PSSLPGPASSPMP:PNSSPL----- 1164
 QY 116 LHTALSSSSSYREAVFKYLSNKESLDQ:VATPGSSDDPIALGVLDKDLFSVPADPNML 175
 DB 1165 -----ASVSVSTSVPLSSSLPI-SVPTTLPAASAPLTIPIAPL----- 1204
 QY 176 DTLVPAHPALVNAIVLHVSAGSAPMPTGDTSSSRMPSSTSYRDMPCGFLFGLDDEDD 235
 DB 1205 -TVSASGPAULTSVTPPLAPVPAAPCPSPSLQSPASPSASALT-GLATAPSLSSSTP 1262
 QY 236 FHP-----NTRSTPSSS---TPSSRPASLYSGAAGRPPTQS----- 270
 DB 1263 GHPULLAP-SSHVPGLNSTVAPACSPVLPASALAS-PFPSAPNPAPAOASLAPASSAS 1321
 QY 271 -ELATALALASTPE-----SSSHPTPTGQGHSS-GT 300
 DB 1322 QALATCAPMAAOTAILAPSAPP-LPLVLAPSPCAAPVLASSQTVPVWAPSTPCT 1381
 QY 301 SPMS-SVQSGTPTINDLSQALQHALQAGQPS-----LQSQMOPQL-----QQLR 346
 DB 1382 SLASAPVAPPTVLAAPSSSTQTLPAVPSPPLPSPASTOTLALAPALPTLGGSPSOTL 1441
 QY 347 DMGLQDCE-----LSLRPCR-----PVTGTSKOPWSSSLLEEHLEPA 384

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 12, 2003, 06:36:13 / Search time 41 Seconds
Without alignments
403 429 Million cell updates/sec

Title: US-09-735-251-4

Perfect score: 2133

Sequence: 1 YSLSDWHAVKADQPLTPY SFGYGVWALXVPSLFV 416

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 4231558 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA*

- 1: /cgn2_6/ptodata/1/1aa/5A_CWRB.pep*
- 2: /cgn2_6/ptodata/1/1aa/5B_CWRB.pep*
- 3: /cgn2_6/ptodata/1/1aa/6A_CWRB.pep*
- 4: /cgn2_6/ptodata/1/1aa/6B_CWRB.pep*
- 5: /cgn2_6/ptodata/1/1aa/PTCS_CWRB.pep*
- 6: /cgn2_6/ptodata/1/1aa/backfiles.pep*

Pred. No. is the number of residues predicted by domain 1. Have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2133	100.0	416	US-09-735-251-4	Sequence 1, Appl
2	2129	99.8	416	US-09-735-251-4	Sequence 2, Appl
3	139.5	6.3	1391	US-09-735-251-4	Sequence 3, Appl
4	139.5	6.3	1391	US-09-735-251-4	Sequence 4, Appl
5	126	5.9	194	US-09-735-251-4	Sequence 5, Appl
6	126	5.9	2972	US-09-735-251-4	Sequence 6, Appl
7	126	5.9	3118	US-09-735-251-4	Sequence 7, Appl
8	122.5	5.7	816	US-09-246-250-12	Sequence 1, Appl
9	122.5	5.7	1187	US-09-246-250-12	Sequence 12, Appl
10	122.5	5.7	1187	US-09-246-250-12	Sequence 28, Appl
11	122.5	5.7	1187	US-09-246-250-12	Sequence 28, Appl
12	122.5	5.7	1210	US-09-246-250-12	Sequence 28, Appl
13	122.5	5.7	1210	US-09-246-250-12	Sequence 28, Appl
14	122.5	5.7	1210	US-09-246-250-12	Sequence 28, Appl
15	121	5.7	1060	US-09-246-250-12	Sequence 2, Appl
16	118	5.5	345	US-09-719-715-2	Sequence 2, Appl
17	118	5.5	345	US-09-719-715-2	Sequence 2, Appl
18	118	5.5	878	US-09-556-706-2	Sequence 2, Appl
19	117	5.5	543	US-09-535-008-63	Sequence 63, Appl
20	117	5.5	577	US-09-535-008-61	Sequence 61, Appl
21	117	5.5	1229	US-09-043-968A-2	Sequence 2, Appl
22	117	5.5	1646	US-09-545-008-67	Sequence 67, Appl
23	117	5.5	1647	US-09-545-008-67	Sequence 67, Appl
24	117	5.5	1649	US-09-545-008-65	Sequence 65, Appl
25	117	5.5	1650	US-09-545-008-65	Sequence 65, Appl
26	117	5.5	1678	US-09-545-008-69	Sequence 69, Appl
27	117	5.5	1679	US-09-545-008-66	Sequence 66, Appl

28	117	5.5	1681	4	US-09-535-008-77	Sequence 77, Appl
29	117	5.5	1682	4	US-09-535-008-73	Sequence 73, Appl
30	116.5	5.5	706	2	US-08-074-967-2	Sequence 2, Appl
31	116.5	5.5	706	2	US-08-553-541B-2	Sequence 2, Appl
32	116.5	5.4	706	3	US-09-268-202-2	Sequence 2, Appl
33	116.5	5.4	706	5	PCT-US94-06669-2	Sequence 2, Appl
34	115.5	5.4	2414	1	US-08-227-546-2	Sequence 2, Appl
35	115.5	5.4	2414	5	PCT-US95-04682-2	Sequence 2, Appl
36	115	5.4	907	3	US-08-783-774-2	Sequence 1, Appl
37	115	5.4	907	4	US-09-328-599A-1	Sequence 1, Appl
38	114.5	5.4	907	5	PCT-US95-04611A-19	Sequence 19, Appl
39	114.5	5.4	1142	2	US-08-993-118-7	Sequence 7, Appl
40	114.5	5.4	1142	3	US-08-845-528C-7	Sequence 7, Appl
41	114.5	5.4	1142	4	US-09-066-281B-7	Sequence 7, Appl
42	114	5.3	1402	4	US-09-125-635-12	Sequence 12, Appl
43	113.5	5.3	2442	4	US-09-514-247A-10	Sequence 10, Appl
44	113	5.3	749	4	US-08-997-685A-10	Sequence 10, Appl
45	113	5.3	1321	2	US-08-317-310A-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-08-725-758A-4
Sequence 4, Application US/08725758A
Patent No. 6160109
GENERAL INFORMATION:
APPLICANT: Reed, Guy
APPLICANT: Client, Christophe Y.
TITLE OF INVENTION: NOVEL PLATELET ACTIVATION PROTEIN
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110 2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/725,758A
FILING DATE: 04-OCT-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 60/005,074
FILING DATE: 4 OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 05433/C20001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-6906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 410 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-725-758A-4

Query Match 100.0% Score 2133, DB 3, Length 410;
Best Local Similarity 100.0%; Pred. No. 4.6e-181;
Matches 410; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY : MSLS-WH-AVK-ADQP-TPKSLRLPETELGYSJGGYSISFLKOLIAKGLQESVDPDEL 60
|||||

```

Db 1 MSLSDMHLAVKADQPLTPKSIILRLPETELGEYSLSGGYSISFLKLIAGKLCESVDPDEL 60
Qy 61 IDLIYCGRLKDDQTLDFYGIQFGSTVHVLRKSWPEPDQKPEVDKVAAMREFRVLTAL 120
Db 1 IDLIYCGRLKDDQTLDFYGIQFGSTVHVLRKSWPEPDQKPEVDKVAAMREFRVLTAL 120
Qy 121 HSSSYREAVEMKLSKESLDQ::VATPGUSSDP::ALGVQDOLFSVADPNMGLTLPV 180
Db 121 HSSSYREAVEMKLSKESLDQ::VATPGUSSDP::ALGVQDOLFSVADPNMGLTLPV 180
Qy 181 AHPALVNAIVLVHVSAGSAPMPGTSSSSRSMSSSSSYROMPGGFLFEGLSDDDDHPNT 240
Db 181 AHPALVNAIVLVHVSAGSAPMPGTSSSSRSMSSSSSYROMPGGFLFEGLSDDDDHPNT 240
Qy 241 RSTPSSSTRSRPASLGYSAGAGPRPTQSELATALALASTPSSSHTPTGTQGHSSGT 300
Db 241 RSTPSSSTRSRPASLGYSAGAGPRPTQSELATALALASTPSSSHTPTGTQGHSSGT 300
Qy 301 SPMSGVSQSTPTINDLFSQALQHALQASQPSQSOWQPOLQOLDMGIQDDLSLRPC 360
Db 301 SPMSGVSQSTPTINDLFSQALQHALQASQPSQSOWQPOLQOLDMGIQDDLSLRPC 360
Qy 361 RPPVGTSKQPMWSSSLLEPHELPAPEPPASCRGYCPWEALMKVPPSLP 409
Db 361 RPPVGTSKQPMWSSSLLEPHELPAPEPPASCRGYCPWEALMKVPPSLP 409

RESULT 2
US-08-725-758A-2
Sequence 2, Application US/0674575A
Patent No. 6160108
GENERAL INFORMATION:
APPLICANT: Reed, Guy
APPLICANT: Clement, Christophe Y.
TITLE OF INVENTION: NOVEL FLAT-BED ACTIVATION METHOD
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson LLP
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110 2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: DOS
SOFTWARE: EASET, Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/04/01/758A
FILING DATE: 04-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 62/005,014
FILING DATE: 06-OCT-1995
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,813
REFERENCE/DOCKET NUMBER: 25431/2209
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5170
TELEFAX: 617-542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-725-758A-2

```

Query Match 99.84, Score 2129, US 3, Length 426,
Best Local Similarity 100.0%, Field No 1, File 180.

```

Matches 409; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSLSDMHLAVKADQPLTPKSIILRLPETELGEYSLSGGYSISFLKLIAGKLCESVDPDEL 60
Db 1 MSLSDMHLAVKADQPLTPKSIILRLPETELGEYSLSGGYSISFLKLIAGKLCESVDPDEL 60
Qy 61 IDLIYCGRLKDDQTLDFYGIQFGSTVHVLRKSWPEPDQKPEVDKVAAMREFRVLTAL 120
Db 61 IDLIYCGRLKDDQTLDFYGIQFGSTVHVLRKSWPEPDQKPEVDKVAAMREFRVLTAL 120
Qy 121 HSSSYREAVEMKLSKESLDQ::VATPGUSSDP::ALGVQDOLFSVADPNMGLTLPV 180
Db 121 HSSSYREAVEMKLSKESLDQ::VATPGUSSDP::ALGVQDOLFSVADPNMGLTLPV 180
Qy 181 AHPALVNAIVLVHVSAGSAPMPGTSSSSRSMSSSSSYROMPGGFLFEGLSDDDDHPNT 240
Db 181 AHPALVNAIVLVHVSAGSAPMPGTSSSSRSMSSSSSYROMPGGFLFEGLSDDDDHPNT 240
Qy 241 RSTPSSSTRSRPASLGYSAGAGPRPTQSELATALALASTPSSSHTPTGTQGHSSGT 300
Db 241 RSTPSSSTRSRPASLGYSAGAGPRPTQSELATALALASTPSSSHTPTGTQGHSSGT 300
Qy 301 SPMSGVSQSTPTINDLFSQALQHALQASQPSQSOWQPOLQOLDMGIQDDLSLRPC 360
Db 301 SPMSGVSQSTPTINDLFSQALQHALQASQPSQSOWQPOLQOLDMGIQDDLSLRPC 360
Qy 361 RPPVGTSKQPMWSSSLLEPHELPAPEPPASCRGYCPWEALMKVPPSLP 409
Db 361 RPPVGTSKQPMWSSSLLEPHELPAPEPPASCRGYCPWEALMKVPPSLP 409

RESULT 3
US-08-545-860D-55
Sequence 55, Application US/08545860D
Patent No. 6040140
GENERAL INFORMATION:
APPLICANT: Croce, Carlo
APPLICANT: Canaan, Eli
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
ADDRESSEE: No. 60401401s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/545,860D
FILING DATE: 07-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE: 22-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/10930
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/327,392
FILING DATE: 19-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/320,559
FILING DATE: 11-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/662,443

```

FILING DATE: 14-MAY-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/971,094
 FILING DATE: 30-OCT-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/886,839
 FILING DATE: 27-MAY-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/805,091
 FILING DATE: 11-DEC-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: DeLuca Esq., Mark
 REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: TCU 1262
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1093 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEetical: NO
 ANTI-SENSE: NO

Query Match 6.5%; Score 139.5; DB 3; Length 1093;
 Best Local Similarity 21.4%; Pred. No. 0.0014;
 Matches 108; Conservative 48; Mismatches 181; Indels 167; Gaps 20;

QY	11	KIADQPLT---	PKSLRLPETELGYSYSGGYS:SEKQOLIAG	-----	KJESVP 56
DB	590	RLSRSPFTSLPSSAS:STTQV--	FSLAGSTFSLPSTHIFGTMGAVNPALLSQAESST	647	
QY	57	DEELIDLVCGRKDKDDTDFYVQPGSTVHV	-----	LRKSWPEPDQKE 102	
DB	648	EPDLEDCFRGRTGTSPOESLS--	SVSPISLFPALPDQASAPCGGQDPAAPGTTNWEQ	705	
QY	103	PVDK-----	VAAMREFVLHTALHSSSYSEAVFKMLSKESLDQIIIVATPGLSSDPI	155	
DB	706	LLEKQGDGEAGNIVVEY:KALHA:CKENQ:R:GEQ:LS:TKAKKERLQILNV	-----	755	
QY	156	ALGVLCQKDLFSVADPNVLDLVPALPAVNAIV:VLHVSVAGSAPMPCD	-----	SSSRSM 212	
DB	756	-----	-----	-----	
QY	213	PSSSYRMYEGGFLEGLSDDEDFHNTSTPSSSTPSSRPASLGYSAGGPRPITQSEL	272		
DB	798	PKSS:G-----	LNLSLTSSD-----	PHSCPSRSSSLSFHTPTPLPQQSP	844
QY	273	AT-ALALASTPSSSHTP-----	TPGTQGHSSGTSRPMSSGVSGTPTINDLFSQALQHA	325	
DB	845	ATPLALGAPAPLPPOQNGLGRAPGAAG--	LGAMPNABGLGLAGSGGLPGLNGLGG	902	
QY	326	L-----	QASQPSLQ-----	SOMQPOLQQLRDMQIDDEL	357
DB	903	LNGAAMPNAPASLSQAGGAPTLCPLFCNLSLTETDORHLQOOQEQLOQLO	-----	QLLA 955	
QY	358	RPCRP-----	VTSKQPSSSSLDEE-----	PHLEPA 384	
DB	956	SPQLTPEHTVYVQMIQ:QCKRCELCOMAGJSLPMASLANSSTPL:SNQTPGLPT	10:5		
QY	385	SPEPPASCRGYCPWALXKVPBS: 408			
DB	1016	ASAPP-----	LUPAGAL--VAPS: 1042		

RESULT 4
 PCT-US94-04496-55
 Sequence 55, Application PC/TUS9404496
 GENERAL INFORMATION:

APPLICANT: Croce, Carlo
 APPLICANT: Canaan, Eli
 TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
 TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
 TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
 NUMBER OF SEQUENCES: 86
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
 ADDRESSEE: Norris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-COS/MS-DOS
 SOFTWARE: Patent in Release #1.0. Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/04496
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: DeLuca Esq., Mark
 REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: TCU-1242
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3100
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 55:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1093 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHEtical: NO
 ANTI-SENSE: NO

Query Match 6.5%; Score 139.5; DB 5; Length 1093;
 Best Local Similarity 21.4%; Pred. No. 0.0014;
 Matches 108; Conservative 48; Mismatches 181; Indels 167; Gaps 20;

QY	11	KIADQPLT---	PKSLRLPETELGYSYSGGYS:SEKQOLIAG	-----	KJESVP 56
DB	590	RLSRSPFTSLPSSAS:STTQV--	FSLAGSTFSLPSTHIFGTMGAVNPALLSQAESST	647	
QY	57	DEELIDLVCGRKDKDDTDFYVQPGSTVHV	-----	LRKSWPEPDQKE 102	
DB	648	EPDLEDCFRGRTGTSPOESLS--	SVSPISLFPALPDQASAPCGGQDPAAPGTTNWEQ	705	
QY	103	PVDK-----	VAAMREFVLHTALHSSSYSEAVFKMLSKESLDQIIIVATPGLSSDPI	155	
DB	706	LLEKQGDGEAGNIVVEY:KALHA:CKENQ:R:GEQ:LS:TKAKKERLQILNV	-----	755	
QY	156	ALGVLCQKDLFSVADPNVLDLVPALPAVNAIV:VLHVSVAGSAPMPCD	-----	SSSRSM 212	
DB	756	-----	-----	-----	
QY	213	PSSSYRMYEGGFLEGLSDDEDFHNTSTPSSSTPSSRPASLGYSAGGPRPITQSEL	272		
DB	798	PKSS:G-----	LNLSLTSSD-----	PHSCPSRSSSLSFHTPTPLPQQSP	844
QY	273	AT-ALALASTPSSSHTP-----	TPGTQGHSSGTSRPMSSGVSGTPTINDLFSQALQHA	325	
DB	845	ATPLALGAPAPLPPOQNGLGRAPGAAG--	LGAMPNABGLGLAGSGGLPGLNGLGG	902	
QY	326	L-----	QASQPSLQ-----	SOMQPOLQQLRDMQIDDEL	357
DB	903	LNGAAMPNAPASLSQAGGAPTLCPLFCNLSLTETDORHLQOOQEQLOQLO	-----	QLLA 955	

```

QY 358 RCRPP-----VCTSKQPMSSSLLE-----PHELPA 384
DB 956 SPCLTEHCTVYQMIOQKRE-CFLQMARSSQLPMASULASSSTRLLASGTGGLST 1015
QY 385 SPEPPASCRGYPWEALMAYVPSL 408
DE 1016 ASAPP-----LIPAGAL -VAPSL 1032

RESULT 5
US-09-735-251-4-19344
; Sequence 19344, Application US/097352991A
; Patent No. 6553795
; GENERAL INFORMATION:
; APPLICANT: Marc C. Rubinfeld et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/7352991A
; PRIOR FILING DATE: 1999-02-13
; PRIOR APPLICATION NUMBER: US 63/034,748
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/734,190
; NUMBER OF SEQ ID NOS: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19344
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-735-251-4-19344

Query Match 5.9%; Score 126; DB 4; Length 394;
Best Local Similarity 22.7%; Pred. No. 0.0343;
Matches 92; Conservative 40; Mismatches 163; Indels 104; Gaps 20;

QY 54 SVDPPELDILVCGRKLDKDDOTLDYGIQ---GSTVHLVLRKSWPPDQKPEPVKVAAMR 121
DB 47 AVHPPDAVALGSAGRL---HGLRFSTQQAHLIRHVVHPDPSLPFGKGPAA 95
QY 112 ERLVHLTAHSSSSYEAVFKMLSNKESLDQIVATPG---SSDPIALGVLDKDLFSVFADP 167
DB 96 - - - - -RETVLRVPPGACGFAARRPPGAFERGTTPVSPGSLERKPKAG 140
QY 168 VFADPNKDLTVFAH-----ALVNAIVLVLSVACSAHPTGD 206
DB 141 CLAPPQLADR PAHPPGSSKLLKAGTHVQKNSLPAHPPGAGHFAHNSRLPCHA 199
QY 207 SSS---PSVPSA-VLHPPDPEFADL---HNNVQVPSSTLHPPASLHDA 261
DB 200 AAARHTRHCAVPPALSKRPAA-----HPPHPPGAGHFAHNSRLPCHA 243
QY 262 AGPPPTQSELATALALASTESSSHHTPTGQKLES---STSPMSSTVSGSGTPTND 316
DB 244 AGRP GREPTAHEFADPATPANGRCPLPACGPAAPGCPGSSNAVTPGHRBAAR 302
QY 317 LFSQALQALQASGOPS-LOSQWQPOLQRLNGQDDHSLRP---CRPVGTSTQOP 370
DB 303 LF-----RALGLPSRYCEVTPHPPHVVAFG---GSKLSPTGAGHARCAVG 349
QY 371 WSSSLLEEPHELPAAP-EPFASCRGYPWEALMAYVPSL 409
DB 350 WRTAAARPR-----SPGRNPPGAGTGRTGFFHMRTPQSP 386

RESULT 6
US-09-579-181-2
; Sequence 2, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chiviva, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247

```

```

; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2972
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-2

Query Match 5.9%; Score 126; DB 4; Length 2972;
Best Local Similarity 22.7%; Pred. No. 0.1; 178; Indels 162; Gaps 23;
Matches 114; Conservative 49; Mismatches 49;

QY 12 LADQPLTPKSL-----RLPETELGEYSLGGYSIFLKQIAGKLOESVDP 58
DB 1068 LAPAPRPSSGLPAVLNPRPTLTGRLPTTGTARAMPPTTLVRLL--KLHVS-PSP 1124
QY 59 ELIDLIYCGRKLDKDDOTLDYGIQ---GSTVHLVLRKSWPPDQKPEPVKVAAMR 115
DB 1125 EV-----SASAPGAAPLTSSPLHV-PSLPGPASSPMPINKSSPL----- 1164
QY 116 LHTALESSSSYEAVFKMLSNKESLDQIVATPG---SSDPIALGVLDKDLFSVFADP 175
DB 1165 -----ASPVSTVSVPSSSLPI-SVPTTLPAASAPLTIP:SAPL----- 1204
QY 176 DTLVPAHPALVNAIVLVLSVACSAHPTGDSSSSSMPSSSSYRDMPGGFLFEGLSDDDD 235
DB 1205 -TVSASGPAALSVTPPLAPVPAAPGFPSPLOSPGASPSASALT-GLATAPSLSSSTP 1262
QY 236 FHP-----NTRSTPSS---TPSSSRPASLVSGAAGPRP:TQS----- 270
DB 1263 GHPLLLAPSSHVPGUNSTVAPACSPVLVPASALAS-PPSPAPNAPAAQASLLAPASAS 1321
QY 271 -EATATALASTPE-----SSSHTPTPTGQHSS-GT 300
DB 1322 QALATLPAKAAPOTAALAPSPAPPALPLVLAPSGAAPVLASSOTVPVVPASSTPT 1381
QY 301 SPMS-SGVSGTPTINDLFSQALQALQASGOPS-----LOSQWQPOL----- 346
DB 1382 SLASAGVFAPTPVLAPSTOTMLPAPVPSPLPSPASTOTLALAPALAPTLCGSSPSOTL 1441
QY 347 DMGIQDDE-----LSLRPCR-----PPVGTSKQPMSSSLLEEPHELPA 384
DB 1442 SLGTGNGQFFPTQ:LSLTPASSLVPTPAQT:SLAPGPLGPTQ---TLSLAPAPLAPA 1498
QY 185 SPEPPASCRGYPWEALMAYVPSL 409
DB 1499 SPVQPA-----PAHTTTLAPAS 1515

RESULT 7
US-09-579-181-1
; Sequence 1, Application US/09579181
; Patent No. 6365372
; GENERAL INFORMATION:
; APPLICANT: Chiviva, John
; APPLICANT: Yaciuk, Peter
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)
; FILE REFERENCE: 16153-4247
; CURRENT APPLICATION NUMBER: US/09/579,181
; CURRENT FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/136,620
; PRIOR FILING DATE: 1999-05-27
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3118
; TYPE: PRT
; ORGANISM: Human
US-09-579-181-1

```


Query Match 5.7% Score 122.5; DB 4; Length 816;
 Best Local Similarity 22.7%; Pred No. 0.11;
 Matches 114; Conservative 49; Mismatches 178; Indels 142; Gaps 23;

DB 12 LACDPLTKSL-----RLPTEIIEYSLSGYSIFLKLGLAKLQESVDP 58
 1214 LAPARPSSGLPAVLNRPPTLTHURLETPFLICARAPKPTLVRLP LKHS-PS 1270
 QY 59 ELIDLIYGRKLKQDQTLFYQIQ - GSTVFLKSKWPEPDKFEVDVNAAMREFV 115
 DB 1271 EV-----SASAPCAAPLTSSILNV PSEDEGPASSEMPIEASSPL 1310
 QY 116 LHTALHSSSYREAVFKLXKESLQOIVATILSSDPLALCVLCKELSVFADNKL 175
 DB 1311 -----ASPVSTSVLSLSLPT-SVPTTLTAFASAPLTIPASAP 1350
 QY 176 DTLVPAHPALNAIVLVHSAVAPMGTSSSRSPSSSYKMPJ3FLPEGLSDPED 235
 DB 1351 -TVSASGPAALLTSVTPPIAPVVFAPAPSPS-OPKCASESASALL-GLATAPS-SSSOTF 1408
 QY 236 FHP-----NTRSFSSS TPSPASLQYSGAAGRPITQS----- 270
 DB 1409 GHPLLAPTSSHPVGLNFTVAFACSEVLVPASALAS-PPPSAENPAPACASALLAPASSA 1467
 QY 271 -ELATALLASTPE-----SSHTPTTRTQCHSS GT 300
 DB 1469 QALATPLAPMAAPQALAFAPAPPLAPLVNAPSPAAVLSQTPVVAISS-PTG 1527
 QY 301 SPMS-SGVOSQTPITNLFSCALQALCASQPS-----LCSQKPGQ-----QQL 346
 DB 1528 SLASAPVAPTPVAPASTQMLFAPVPSLHPASTOTLALAPALPTGSSPQTL 1587
 QY 347 DMGICQDE-----LSLRPP-----FPVTSKQPASSSLEDEPHLEDA 164
 DB 1588 SLGTGNGCPPTQTLSTPASASVETPAQTLKAPGPPGPTG TLSAPAPPLAPA 1644
 QY 385 SPERPASTRGYCPMEALVKVPS 407
 DB 1645 SPVGFQ-----PAHTLTAPAS 1641

RESULT 8
 US-09-266-225D-12
 : Sequence 12, Application US/09266225D
 : Patent No. 6573364
 : GENERAL INFORMATION:
 : APPLICANT: Nandabalan, Krishan
 : APPLICANT: Kirsstore, Stephen
 : APPLICANT: TCGenerex, Veriza
 : TITLE OF INVENTION: Isolation and characterization of human and monkey indk
 : TITLE OF INVENTION: Syndrome (HIS) Protein Defects and HES Protein
 : TITLE OF INVENTION: Interacting Proteins
 : FILE REFERENCE: 15966-523
 : CURRENT APPLICATION NUMBER: US/09266-225D
 : CURRENT FILING DATE: 1999-03-10
 : NUMBER OF SEQ ID NOS: 19
 : SOFTWARE: Patent In Ver. 2.1
 : SEQ ID NO 12
 : LENGTH: 816
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-266-225D-12

Query Match 5.7% Score 122.5; DB 4; Length 816;
 Best Local Similarity 25.1%; Pred No. 0.01;
 Matches 85; Conservative 39; Mismatches 116; Indels 99; Gaps 18;

QY 90 LRKSWPEPDKPEVDKVA-----NREFVLHTALHSSSSSYREAVFKLSKESLQIIVA 146
 DB 491 LRDPGSAPLEAPKPKVPTACERQEREKPRQERAKERE---KRPQERERKRCAGA 547
 QY 147 TPGLSDDPIALGVLDKDLDFSVFADPNMLDTLV-----PAHPALVNAIVLVHSAVAGAPM 202

DB 548 SGQSTDPHAGLVLSN-----DRSLERWTRMARPAAPALT-----SVPAPAPA 592
 QY 203 PGTDSSSRSSVSSYRDMPCGFLPEGLSDDEDDFHENTRTSTPSSSTPSSRPA-----SL 256
 DB 593 P-TPPTPVPQPTSP-----PSGVAQPTG-----PQPSAGSTSGVPQCPACPPGPAP 640
 QY 257 GYSGAAGPPI-TQSELATALALASTPSSSHPTPTGTQCHS-----SG 299
 DB 641 HPTSPGPIIVPAPPQTATSTSLA---AOSLVPPPLPGSSTPGVLVPPFPGCLPPPDAG 697
 QY 300 TSMSSSOVQSGSTP-----ITNDFLSQALQHALQA--SCOP-----SLOSQKQPLQOLR 346
 DB 698 GACQSS--XSESQVNVLTQOLSKSQVEDPLPVFSGTPKSGAGYGVGPDLEFLNQSF 755
 QY 347 DMGICQD-----DELSLRCPRPVGTSGKQPMSSSLLEE 378
 DB 756 DMGVADGPGDQADASL-----SASLLAD 780

RESULT 9
 US-08-320-559-28
 : Sequence 28, Application US/08320559
 : Patent No. 5633135
 : GENERAL INFORMATION:
 : APPLICANT: Croce, Carlo
 : APPLICANT: Canaan, Eli
 : TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
 : TITLE OF INVENTION: Detection and Treatment of Acute Leukemias
 : TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
 : TITLE OF INVENTION: All-1 Region
 : NUMBER OF SEQUENCES: 44
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5633135ris
 : STREET: One Liberty Place - 46th Floor
 : CITY: Philadelphia
 : STATE: PA
 : COUNTRY: USA
 : ZIP: 19103
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
 : COMPUTER: IBM PS/2
 : OPERATING SYSTEM: PC-DOS
 : SOFTWARE: WORDPERFECT 5.1
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/320,559
 : FILING DATE:
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/062,443
 : FILING DATE: 14 MAY 1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/971,094
 : FILING DATE: 30-OCT-92
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/888,830
 : FILING DATE: 27-MAY-92
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/07/805,093
 : FILING DATE: 11-DEC-91
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Deluca, Mark
 : REGISTRATION NUMBER: 33,229
 : REFERENCE/DOCKET NUMBER: TJU-0855
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (215) 568-3100
 : TELEFAX: (215) 568-3439
 : INFORMATION FOR SEQ ID NO: 28:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1187 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein

US-08-320-559-28

Query Match 5.7%; Score 122.5; DB 3; Length 1187;
 Best Local Similarity 21.1%; Pred. No. 0.053;
 Matches 88; Conservative 52; Mismatches 171; Indels 107; Gaps 18;

QY 8 LAVKACQPLTEKSIIL----RJETELGEYSIGCYISIFLKQLIA-----GKQESVPD 57
 DB 226 LAVKHKKE-TPQSLVAPAQPSQTFPPSPKSVAMQCKPTAYVRPMQCCQARSES 284
 QY 58 PELICLIYGRKKDD---QTLDFYGIQPGSTVHLRKSWPEPCQPEPVKVAAMREF- 113
 DB 285 PEL-----KPLPEDYRQOTTEKTLKVPAAKUTLKNPQSQVEQYTSNEVHCVEIL 337
 QY 114 -RVLR-----TALHSSSYREAVFQMLSNKESLQIIVATGLSSDIALGVLCQDLF 166
 DB 338 KEMTHSNPPPTTAHTTESTAEPSKEP FPKLS-----NSQGTSSMLEDDLQ----- 405
 QY 167 SVFADPNMLDTLVFAHPALYNALVHWRVAGSAFVPGTDSRSSRMPSSSYEDKPGGFLF 226
 DB 374 SVTONKQCYDTSSKTHS-----NSQGTSSMLEDDLQ----- 405
 QY 227 EGLSDDED-DPHNTRSTPSSSTPSSPPASLGVSGAAGPRPTQSELATALA-LASTPES 284
 DB 406 --LSDSDSDSEQTPKEKPPSSAPPAPQSL-----PEPVASAHSSASESTSDSDS 456
 QY 285 SSHTPTPTQCHSSGTSPMSSGVQSGTPTINDLF--SOALQHALQAGOPSLQSQWQPOL 342
 DB 457 SSDSESSSSSDSENEPELETPEPEPTTNKWLQCNLTKVQSPAPAPPEPRSTEPPR 516
 QY 343 QQLRDYICQDELRLPCR---PPVGTSGKOPKSSSLLEPH-----ELPASPEPP 389
 DB 517 RHPESKSSSASATSCHESESKDPKPPSSSKAPRAPP--EAPHGKRSCKSPAQCEPP 572

RESULT 11

US-08-545-8600-28
 Sequence 28, Application US/08545860

Patent No. 6043140

GENERAL INFORMATION:

APPLICANT: Ciba, Carle

APPLICANT: Ciba, Carle

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

APPLICATION NUMBER: US 08/320,559
 FILING DATE: 11-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/062,443

FILING DATE: 14-MAY-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/971,394

FILING DATE: 10-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/888,839

FILING DATE: 27-MAY-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/805,092

FILING DATE: 11-DEC-1991

ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Esq., Mark

REGISTRATION NUMBER: 33,229

REFERENCE/DOCKET NUMBER: 133-1262

TELECOMMUNICATION INFORMATION:

TELEPHONE: (215) 568-3100

TELEFAX: (215) 568-3439

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1187 amino acids

TYPE: amino acid

TOPLOGY: linear

MOLECULE TYPE: protein

US-08-545-8600-28

Query Match 5.7%; Score 122.5; DB 3; Length 1187;

Best Local Similarity 21.1%; Pred. No. 0.053;

Matches 88; Conservative 52; Mismatches 171; Indels 107; Gaps 18;

QY 8 LAVKACQPLTEKSIIL----RJETELGEYSIGCYISIFLKQLIA-----GKQESVPD 57

DB 226 LAVKHKKE-TPQSLVAPAQPSQTFPPSPKSVAMQCKPTAYVRPMQCCQARSES 284

QY 58 PELICLIYGRKKDD---QTLDFYGIQPGSTVHLRKSWPEPCQPEPVKVAAMREF- 113

DB 285 PEL-----KPLPEDYRQOTTEKTLKVPAAKUTLKNPQSQVEQYTSNEVHCVEIL 337

QY 114 -RVLR-----TALHSSSYREAVFQMLSNKESLQIIVATGLSSDIALGVLCQDLF 166

DB 338 KEMTHSNPPPTTAHTTESTAEPSKEP FPKLS-----NSQGTSSMLEDDLQ----- 405

QY 167 SVFADPNMLDTLVFAHPALYNALVHWRVAGSAFVPGTDSRSSRMPSSSYEDKPGGFLF 226

DB 374 SVTONKQCYDTSSKTHS-----NSQGTSSMLEDDLQ----- 405

QY 227 EGLSDDED-DPHNTRSTPSSSTPSSPPASLGVSGAAGPRPTQSELATALA-LASTPES 284

DB 406 --LSDSDSDSEQTPKEKPPSSAPPAPQSL-----PEPVASAHSSASESTSDSDS 456

QY 285 SSHTPTPTQCHSSGTSPMSSGVQSGTPTINDLF--SOALQHALQAGOPSLQSQWQPOL 342

DB 457 SSDSESSSSSDSENEPELETPEPEPTTNKWLQCNLTKVQSPAPAPPEPRSTEPPR 516

QY 343 QQLRDYICQDELRLPCR---PPVGTSGKOPKSSSLLEPH-----ELPASPEPP 389

DB 517 RHPESKSSSASATSCHESESKDPKPPSSSKAPRAPP--EAPHGKRSCKSPAQCEPP 572

RESULT 11

PCT-US94-04496-28

Sequence 28, Application PC/TUS9404496

GENERAL INFORMATION:

APPLICANT: Ciba, Carle

APPLICANT: Ciba, Carle

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

TITLE OF INVENTION: For the treatment of acute leukemia

ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
 ADDRESSEE: Norris
 STREET: One Liberty Place, 46th floor
 CITY: Philadelphia
 STATE: Pennsylvania
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/04496
 FILING DATE:
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: DeLuca Esq., Mark
 REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: RU 1042
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3300
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 28
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1187 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 PCT US94-04496-28

Query Match 5.78; Score 122.5; PB 1; Length 1187;
 Best Local Similarity 21.1%; Pred. No. 0.054;
 Matches 88; Conservative 52; Mismatches 171; Indels 107; Gaps 18;
 QY 8 LAVKLADQLTPKSIJ-----RLPETELGEYSIGGVYSISFLKOLIA-----GKLOESVDP 57
 DB 226 LAVKVDKE TPQDSLVAAPQPSQTFPPPSLPSKSVAMQOKPTAYVRPMDGQQQAPSES 307
 QY 58 PELIDLLYGGRKLKDD---QTLDFYGIQPGSTVHVLRKSWPEPDOKPEPVDKVAAMREF- 113
 DB 308 PEL-----KELPDYRQQTFEKDLKVPKAKLTKLKPVSQSVQETYSNEVHCVEIL 360
 QY 114 -RVLH-----TALHSSSYREAVFKVLSNKSLEDOIIVATPGLSSDPALGLVDKDLF 166
 DB 361 KEMTHSWPPPLTAHTPTAETSKPEP-FPTKDS-----QHVS 396
 QY 167 SVFADPNVLDLVPAPALVNAVLVLSHVSAGSAPMPTDSSSRMPSSSYRDMFGGLF 226
 DB 397 SVTQKQGYDTSSKTHS-----NSQGTSSMLLEDQLQ----- 428
 QY 227 EGLSDDED-DHPNTRSTPSSSTPSSRPASLCYSCAAGRPITQSELATALA-LASTPES 284
 DB 429 ---SSSESSCSSEQTPEKPPSSAPSAFQSL-----PEPVASHSSASESTSDS 479
 QY 285 SSGHTFTGTQHSSGTSPMSGVQSGTPTITKDLF--SQALQHALQASQPSLQSQWQPOL 342
 DB 480 SSSSESSSSSSSENEPLETAPPEPEPT--NKWOLDNWLTKVSPAAPPEGPRSTPEPR 539

US-09-735-251-4-28
 US-09-735-251-4-28
 Sequence 26, Application US/08320559
 Patent No. 5633135
 GENERAL INFORMATION:
 APPLICANT: Croce, Carlo
 APPLICANT: Canani, Eli
 TITLE OF INVENTION: Diagnostics, Therapeutics and Methods for
 TITLE OF INVENTION: Detection and Treatment of Acute Leukemias

TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the
 TITLE OF INVENTION: All-1 Region
 NUMBER OF SEQUENCES: 44
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5633135ris
 STREET: One Liberty Place - 46th Floor
 CITY: Philadelphia
 STATE: PA
 COUNTRY: USA
 ZIP: 19103
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
 COMPUTER: IBM PS/2
 OPERATING SYSTEM: PC-DOS
 SOFTWARE: MICROPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/320,559
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/062,443
 FILING DATE: 14 MAY 1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/971,094
 FILING DATE: 30-OCT-92
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/898,830
 FILING DATE: 27-MAY-92
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/07/805,093
 FILING DATE: 11-DEC-91
 ATTORNEY/AGENT INFORMATION:
 NAME: DeLuca, Mark
 REGISTRATION NUMBER: 33,229
 REFERENCE/DOCKET NUMBER: 707-0855
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (215) 568-3300
 TELEFAX: (215) 568-3439
 INFORMATION FOR SEQ ID NO: 26:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1210 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-320-559-26

Query Match 5.78; Score 122.5; PB 1; Length 1210;
 Best Local Similarity 21.1%; Pred. No. 0.054;
 Matches 88; Conservative 52; Mismatches 171; Indels 107; Gaps 18;
 QY 8 LAVKLADQLTPKSIJ-----RLPETELGEYSIGGVYSISFLKOLIA-----GKLOESVDP 57
 DB 249 LAVKVDKE TPQDSLVAAPQPSQTFPPPSLPSKSVAMQOKPTAYVRPMDGQQQAPSES 307
 QY 58 PELIDLLYGGRKLKDD---QTLDFYGIQPGSTVHVLRKSWPEPDOKPEPVDKVAAMREF- 113
 DB 308 PEL-----KELPDYRQQTFEKDLKVPKAKLTKLKPVSQSVQETYSNEVHCVEIL 360
 QY 114 -RVLH-----TALHSSSYREAVFKVLSNKSLEDOIIVATPGLSSDPALGLVDKDLF 166
 DB 361 KEMTHSWPPPLTAHTPTAETSKPEP-FPTKDS-----QHVS 396
 QY 167 SVFADPNVLDLVPAPALVNAVLVLSHVSAGSAPMPTDSSSRMPSSSYRDMFGGLF 226
 DB 397 SVTQKQGYDTSSKTHS-----NSQGTSSMLLEDQLQ----- 428
 QY 227 EGLSDDED-DHPNTRSTPSSSTPSSRPASLCYSCAAGRPITQSELATALA-LASTPES 284
 DB 429 ---SSSESSCSSEQTPEKPPSSAPSAFQSL-----PEPVASHSSASESTSDS 479
 QY 285 SSGHTFTGTQHSSGTSPMSGVQSGTPTITKDLF--SQALQHALQASQPSLQSQWQPOL 342
 DB 480 SSSSESSSSSSSENEPLETAPPEPEPT--NKWOLDNWLTKVSPAAPPEGPRSTPEPR 539

QY 343 QQLRDMG:QDDHSLRPCR-----PVTGTSKQWSSLSLEERH-----ELFASPEPP 349
DB 540 RHPEKSGSDSATSQEHSESKOPPPSSSKAPRAPP--EAPHGKRSQKSPAQOEPP 595

RESULT 13
US-08-545-8600-26
Sequence 26, Application US/0845458600
Patent No. 6040140
GENERAL INFORMATION:
APPLICANT: Crocco, Carlo
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1 Region
NUMBER OF SEQUENCES: 94
CORRESPONDENCE ADDRESS: 94
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/0845458600
FILING DATE: 07-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE: 22-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE: 09-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 19 OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 19 OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 14-MAY-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/971,094
FILING DATE: 30 OCT-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/989,849
FILING DATE: 27-MAY-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/989,849
FILING DATE: 11-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU 1262
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
TOPOLOGY: linear
X-MOLECULE TYPE: Protein
US-08-545-8600-26

Query Match 5.71; Score 122.5; DB 3; Length 1210;
Best Local Similarity 21.11; Pred. No. 0.054;
Matches 89; Conservative 52; Mismatches 171; Indels 107; Gaps 18;

QY 8 LAVKLADQELTPKSIIL-----RUPETELGEYSUGYGISIFLKOLIA-----GKQESVVD 57
DB 249 LAVKVKHKE TPQDSLVAQAQPPSQTFPPPPSPKSVAMQOKPTAYVVRPMDGQQAQPS 307
QY 58 PELLDLYGKRLKDD---QTLDYFCIQGISTVHVLRKSWPEPCKPEPVCKVAMREF- 113
DB 308 PEL---KELPEYRQCTTEKTDKVPKAKUKLKMPSQSVQETYSKEVHCVEIL 360
QY 114 -RVLH-----TALHSSSYREAFVFMKLSKESLDQI:VATPGLSSDPALGV:QDKDLF 166
DB 361 KEMTHSWPEPTAIHTPTAEPKFP-PPTKOS-----QHVS 396
QY 167 SVFACFNMILFLVPAIPALNAIVLVHSAVAGAMPCTDSSSRMPSSSYRDMPGGLF 226
DB 397 SVTONQVQYDTSSKTHS-----NSCQGTSSMLEDDIQ----- 428
QY 227 EGLSDDED:DFHPNTRSTPSSSPASLGYSGAAGRPITOSELATALA--LASTPES 284
DB 429 --LSDSEISDSEQTPEKPPSSAPPSPAPSL-----PEPVASAKSSSAESESTSDSDS 479
QY 285 SSHTPTRTTQGHSSGTSPMSSGVQSCFTPIITCLF--SOALCHALCASQPSLOSOMQOL 342
DB 480 SSSSESSSSSSDSENEPLETAPPEPPTTKWQDNLWLTKVSPAAPPEGPRSTPEPP 539
QY 343 QQLRDMG:QDDHSLRPCR-----PVTGTSKQWSSLSLEERH-----ELFASPEPP 349
DB 540 RHPEKSGSDSATSQEHSESKOPPPSSSKAPRAPP--EAPHGKRSQKSPAQOEPP 595

RESULT 14
PCT-US94-04496-26
Sequence 26, Application PCT/US9404496
GENERAL INFORMATION:
APPLICANT: Crocco, Carlo
TITLE OF INVENTION: Diagnostics, Therapeutics and Methods
TITLE OF INVENTION: for Detection and Treatment of Acute Leukemias
TITLE OF INVENTION: Resulting from Chromosome Abnormalities in the All-1
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz &
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04496
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeLuca Esq., Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: TJU-1242
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 1210 amino acids
TYPE: amino acid
TOPOLOGY: linear
X-MOLECULE TYPE: Protein
PCT-US94-04496-26

PCT US94-04496-26

Query Match 5.7%; Score 122.5; DB 5; Length 1210;
Best Local Similarity 21.1%; Pred. No. 0.054;
Matches 88; Conservative 52; Mismatches 171; Indels 107; Gaps 18;

QY 8 LAKVLADQPTPKSL-----RLEPTELEGEYSLGYSISLKLQIA-----GKLOESVVD 57
DB 249 LAKVKVDKE-TPCDSLVAPAQPPGUTTFPTFS-PSKSVAMQCKPTAYVPRVGGCQAPSES 307
QY 58 PELIDLIYGRKIKD---CTLDKFGIOPGTVVHLRKSWEPPDOCKPKNKVAAYREF- 113
DB 308 PEL-----KPLPDYRQCTFERDLKVPALAKTKLMKSCSEVQYEVHCVFEL 360
QY 114 -RVH-----TALHSSSSSYREAVFKYLSNKSLSQTVATFGLSDFIALGVLOOKLE 166
DB 361 KEMTWSRPPLTAIHTPTSAESYEP-PPTKUS-----QHVS 396
QY 167 SVFADPNMCLTLVPAHPALVNAIVLVHSAVSAPMPGCTOSSSRMPSSSYECMPGGFLP 226
DB 397 SVTQNKQYDTSSKTHS-----NSQQTSSMLEDLQ----- 428
QY 227 EGLSDDED-DFHPTNSTSSSTSPSPASLSYGAAGPPPTTQSELATALA-LASTPES 284
DB 429 --LSSEDCSCSECTEKPESSSAHSAPSL-----PEPVASASSGASSESTSDUS 479
QY 285 SSGHTPTQTGHSSSTSPMSSGVGCTPTNLF SQALQHALQASCPESLQSQWQQL 342
DB 480 SSSSESSSSSSSENEPLETANUPPETTKKQDLNALTQVSGPAAPEGPRSTEP 539
QY 343 QQLRMGIGDDELSRFR-----FVUTSKQPKSSLSLEPH-----ELPASPEPP 382
DB 540 RHPEKSGSCSATSQESHSESKOPFKSSKAPRAFF-EPHPKESQKSPAQOEP 595

RESULT 15

US-09-911-393-2

Sequence 2, Application US/289-11393

Patent No. 6323008

GENERAL INFORMATION:

APPLICANT: BELLETIER, Marc

APPLICANT: PARKER, William A.

APPLICANT: HAKES, David C.

APPLICANT: ZOFF, David A.

TITLE OF INVENTION: METHODS FOR PRODUCING

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & REMYUS LLP

STREET: 1155 Avenue of the Arts, 4F

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/911,393

FILING DATE: 14-AUG-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7189-312-999

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212)7909090

TELEFAX: (212)8699741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1060 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-911-393-2

Query Match 5.7%; Score 121; DB 4; Length 1060;

Best Local Similarity 22.1%; Pred. No. 0.06;

Matches 101; Conservative 55; Mismatches 194; Indels 108; Gaps 21;

QY 6 WHLAVKIAD-----QPLTPKSLRLPE---TELGEYSLGYSIS----- 41
DB 577 YHVVLTMANKIGSVVIDCEPLEGGQTVVPDGRTPDISHFYVGGYGRSDMPT-SHVTYNN 636
QY 42 ---FLKCLIAKLCESVDPDELIDLIYCGRKLDQDQDFYGIOPGTVVHLRKSWEPPD 98
DB 637 VLLYRNCLNAEEIRTLFSLQDLIG-----TEAHMGSSSSSAH-----STPS-- 678
QY 99 OKPEPVEKVAAYREFRVLHTALHSSSSSYREAVFKYLSNKSLSQTVATP-----GLSS 152
DB 678 ---TPALNGAHS-TPTPADSSAHSPTS-----TPADSSAHS-TSPADNGAHS 723
QY 153 DPALGVCKDKJESVPADPNMCLTLVPAHPALVNAIVLVHVS-----AGSAPMPG 204
DB 724 TPSTPGSSAHS-TPTPADNGAHS-T--PSAPASNA-----HSTPSTPADNGAHS-TPT 776
QY 205 TDSRSSRSESS-----SYRDMFGGLFEGCL-SDDEDDFHPNTRSTPSSSTPSSRPSL 256
DB 777 ADNGAHS-TPTPGDNGAHS-TPTPGDSSAHS-TTPTPADNGAHS-TSPADSNNAHS-TPT 836
QY 257 GYSGAAGPPPTQSELATALALASTP-ESSSH-TPT-PGTQG-HSSGTSPMSSGVQSGTPT 312
DB 837 GONGAHS-TPSAPAD-SNAHS-TPTPADSSAHS-TSPADNGAHS-TSPADSSAHSPTS 894
QY 313 ITNDLFSCALQHALQASGOPSLSQSQWQQLQOOLRDMGIQDELRLRCRPPVGTSKPMS 372
DB 895 APGONGAHS-TSPADNGAHSPTS-----APGDSNAHS-TPTPADSSAHSPTS 942
QY 373 SSLIEEHHELPAEP-EPASCRGYCPWEALMKVPSPSLP 409
DB 943 TPADSSAHS-TSPADNGAHS-TSPADSSAHSPTS-P 980

Search completed: November 12, 2003, 06:40:02

Job time : 44 secs